CLUSTERING INFLUENCE ON MTPL PREMIUM ESTIMATION USING CREDIBILITY APPROACH

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

İSMAİL ONUR KIZILOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ACTUARIAL SCIENCES

JANUARY 2023

Approval of the thesis:

CLUSTERING INFLUENCE ON MTPL PREMIUM ESTIMATION USING CREDIBILITY APPROACH

submitted by **İSMAİL ONUR KIZILOĞLU** in partial fulfillment of the requirements for the degree of **Master of Science in Actuarial Sciences Department, Middle East Technical University** by,

Prof. Dr. A. Sevtap Kestel
Director, Graduate School of **Applied Mathematics** _______________

Prof. Dr. A. Sevtap Kestel
Head of Department, **Actuarial Sciences** _______________

Prof. Dr. A. Sevtap Kestel
Supervisor, **Actuarial Sciences, METU** _______________

Dr. Bükre Yıldırım Külekci
Co-supervisor, **Mathematics, Technical University of Kaiserslautern** _______________

Examining Committee Members:

Prof. Dr. Ceylan Yozgatlıgil
Statistics Department, METU _______________

Prof. Dr. A. Sevtap Kestel
Actuarial Sciences Department, METU _______________

Assist. Prof. Dr. Başak Bulut Karageyik
Actuarial Sciences Department, Hacettepe University _______________

Date: _______________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: İSMAİL ONUR KIZILOĞLU

Signature :

ABSTRACT

CLUSTERING INFLUENCE ON MTPL PREMIUM ESTIMATION USING CREDIBILITY APPROACH

Kızıloğlu, İsmail Onur
M.S., Department of Actuarial Sciences
Supervisor : Prof. Dr. A. Sevtap Kestel
Co-Supervisor : Dr. Bükre Yıldırım Külekci

January 2023, 95 pages

Pricing is one of the crucial aspects of the insurance industry. There are compulsory regulations on Motor Third Party Liability Insurance in Türkiye. After the cap premium, the significance of correct pricing for the profitability of the companies has increased even more. In this thesis, premiums are estimated quarterly using Bühlmann and Bühlmann-Straub credibility methods for the cities of Türkiye. First, the cities are sorted according to the total number of claims and divided into at least 3 and at most 10 groups. The reason for sorting by the number of claims is that the exposure measure is used for Bühlmann-Straub credibility, and cities that show similarity according to claim number are grouped together. Secondly, credibility results are calculated for cities in which grouping is made by geographical regions. After that, cities are divided into 3, 6, and 9 groups by applying k-means clustering and credibility results are computed for cities. Lastly, the hierarchical method is used for clustering and cities are divided into groups from 2 to 7. As a result of grouping and calculations, it is seen that the Bühlmann-Straub credibility method with 6 groups clustered according to the hierarchical clustering gives better results.

Keywords: Credibility, Bühlmann, Bühlmann-Straub, K-means Clustering, Hierarchical Clustering

ÖZ

KREDİBİLİTE YAKLAŞIMIYLA MTPL PRİM TAHMİNİNDE KÜMELEME ETKİSİ

Kızıloğlu, İsmail Onur
Yüksek Lisans, Aktüerya Bilimleri Bölümü
Tez Yöneticisi : Prof. Dr. A. Sevtap Kestel
Ortak Tez Yöneticisi : Dr. Bükre Yıldırım Külekci

Ocak 2023, 95 sayfa

Fiyatlandırma, sigorta sektörünün en önemli konularından biridir. Türkiye'de Karayolları Motorlu Araçlar Zorunlu Mali Sorumluluk Sigortası için uygulanan zorunlu düzenlemeler bulunmaktadır. Bu düzenlemelerle birlikte gelen tavan prim uygulamasının ardından doğru fiyatlamanın firmaların karlılığı açısından önemi daha da artmıştır. Bu tezde Bühlmann ve Bühlmann-Straub kredibilite yöntemleri kullanılarak Türkiye illeri için çeyrek dönemlik prim tahminleri yapılmıştır. İlk olarak, iller toplam hasar sayısına göre sıralanır ve en az 3, en fazla 10 gruba ayrılır. Hasar sayısına göre sıralamanın nedeni, Bühlmann-Straub güvenilirliği için maruz kalma ölçüsünün kullanılması ve hasar sayısına göre benzerlik gösteren şehirlerin birlikte gruplandırılmasıdır. İkinci olarak, coğrafi bölgelere göre gruplama yapılan iller için kredibilite sonuçları hesaplanmıştır. Daha sonra k-ortalama kümelemesi uygulanarak iller 3, 6 ve 9 gruba ayrılmış ve hesaplamalar yapılmıştır. Son olarak hiyerarşik kümeleme yöntemiyle şehirler 2'den 7'ye kadar gruplanmış ve primler hesaplanmıştır. Bu işlemler sonucunda hiyerarşik kümelemeye göre 6 gruplu Bühlmann-Straub kredibilite yönteminin diğer kümeleme yöntemlerinden daha iyi sonuç verdiği görülmüştür.

Anahtar Kelimeler: Kredibilite, Bühlmann, Bühlmann-Straub, K-ortalama Kümeleme, Hiyerarşik Kümeleme

ACKNOWLEDGMENTS

I would like to express my very great appreciation to my thesis supervisor Prof. Dr. A. Sevtap Kestel and Dr. Bükre Yıldırım Külekci for their patient guidance, enthusiastic encouragement, and valuable advice during the development and preparation of this thesis.

Last but most, I would like to thank my mother for her unwavering support every single day throughout my academic life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

MTPL	Motor Third Party Liability
NCD	No-claim Discount
PPRSA	Private Pension Regulation and Supervision Agency
MOD	Motor Own Damage
TCIP	Turkish Catastrophe Insurance Pool
UW	Underwriting
TL	Turkish Lira
Std. Dev.	Standard Deviation
IAT	Insurance Association of Türkiye
GLM	Generalized Linear Model
CTE	Conditional Tail Expectation
VaR	Value at Risk
CoV	Coefficient of Variation

CHAPTER 1

INTRODUCTION

In the insurance field, two essential factors that have a crucial impact on the financial situation are reserving and pricing. These factors are very interdependent from an actuarial perspective. The situation of getting results such as high loss ratios or high incurred claims or making wrong decisions in one of these factors will very likely affect the other factor directly. It will lead to mistakes in the decisions to be taken for financials. Therefore, in order for insurance companies not to experience any financial difficulties, these two factors should be followed in line and should not be considered separately.

Motor Third Party Liability Insurance (MTPL) is one of the most significant branches for non-life insurance companies in Türkiye. There are many reasons for the significance of the MTPL insurance branch. First of all, it is compulsory insurance and when premium production of insurance companies in Türkiye is examined, the highest written premium in the insurance sector between 2016 and 2021 is from MTPL insurance in accordance with the sector report of the Insurance Association of Türkiye published in 2021 [1]. Moreover, in the same report, another value showing the importance of MTPL insurance for the Turkish insurance sector is that 29 of the 43 non-life insurance companies in Türkiye provide coverage for MTPL insurance. Considering these situations, the importance of following a correct pricing strategy is clearly seen as very critical both for the sustainability of the insurance industry and the economy of Türkiye.

Insurance companies can implement pricing strategies with many methods. They can use these methods alone or create their tariffs by using some methods together. The most crucial features that should be in the methods used for premium prediction are that the results are easy to understand and apply. In light of these features, Bühlmann and Bühlmann-Straub credibility methods are among the methods that can be used for pricing for insurance companies.

In Türkiye, cap premium regulations have been applied for MTPL insurance since April 2017 [5]. Cap premiums are determined by vehicle type, the no-claim discount (NCD), and city. One of the main purposes of the cap premium application is to sell policies at more affordable prices. Secondly, it is to ensure that the right price is reflected in the right customer. However, cap premium applications might impact insurance companies negatively at some points such as selling policies with insufficient premiums to high-risk drivers. Therefore, the importance of premium calculation has risen especially for policies that do not be sold with cap premiums.

Insurance companies can use many applications to guide their pricing strategies with credibility methods. First, companies can estimate their own cap premiums from their own data and create their own risk maps in accordance with cities. The Insurance and Private Pension Regulation and Supervision Agency (PPRSA) announces the cap premiums in Türkiye. Credibility methods can also be applied in this cap premium application, which is calculated by the institution and which all companies are obliged to comply with. One of the most important benefits of using these methods in insurance companies is that they will be able to compare their cap premiums with the premiums announced by the Insurance and Private Pension Regulation and Supervision Agency (PPRSA). By comparison, they can examine where they should comply with the cap premium application or in which cities they can make a profit and write policies at lower prices.

Another advantage of credibility methods is that companies can classify and filter cities according to different factors while making forecasts with these methods. An example of this situation can be given as follows: If the company's data is sufficient, price estimates can be made in cities according to certain vehicles. As a second example, cities can be classified according to the number of claims, and premium estimates can be made with grouped data in this way with credibility. Another conclusion to be derived from here is that credibility methods provide the opportunity to estimate and compare premiums according to various scenarios with the convenience they provide in terms of applicability.

As a result, premium estimation methods can be adapted and improved by insurance companies based on credibility methods. Changing by data filters and groups premium and tariff structures are able to be built and these can allow the dynamic and followable pricing structures for insurance companies. This thesis aims to calculate the average premiums for the cities by grouping the cities in Türkiye with the Bühlmann and Bühlmann-Straub methods and to determine which of the groups and clustering methods used will give the best results for Türkiye.

This thesis is organized as follows: Chapter 2 includes a literature review. In Chapter 3, the features and structure of MTPL insurance in Türkiye are presented. Chapter 4 presents the Bühlmann and Bühlmann-Straub credibility methods that we apply to the MTPL dataset. Chapter 5 includes the k-means and hierarchical clustering algorithms that are used for city grouping in premium calculation. In Chapter 6, descriptive statistics of the dataset, grouping methods, and the comparison of results are presented. Finally, Chapter 7 provides the conclusions of the findings.

CHAPTER 2

LITERATURE REVIEW

From the historical perspective of insurance pricing modeling, the "credibility" term has been about a hundred years past from nowadays. This term was proposed by Mowbray [25] and Whitney [37] for the first time. The first utilization of credibility focused on a limited fluctuation approach. The limited fluctuation approach is based on the calculation of your pricing performance based on your own historical experience and insurance manual.

Mowbray [25] first gave an explanation about how the companies which are large enough can use credibility to model prices exclusively. After then, Whitney [37] proposed the weighting of the total collective experience and individual experience together. This model consists of both the past data that comes from risk and current data at the calculation moment and how to weigh each of them. By this cornerstone, successive studies have aimed to optimize these weights based on different models and distributional assumptions.

In actuarial literature, the credibility premium formula was developed by Waters [35] with a limited fluctuation approach. This model forms the basis of the models used today. After the development of this model, Nelder and Verrall [27] studied the credibility based on the generalized linear model (GLM) approach, this model gave a reasonable solution for premium ratings and reserving problems.

The most widely used model was Bühlmann's credibility, proposed in the 1960s [8], which uses both variances within and between groups. Bühlmann's credibility model was based on the Bayesian approach and distribution assumption for each risk. Bühlmann credibility premium has an important property, that minimizes the expected loss of linear predictor.

Bühlmann and Straub [10], extend the Bühlmann credibility model with different claim experiences of each risk group with different exposures in 1970. Hachemeister [18] proposed a regression credibility model, which can be explained as the covariate to the conditional mean of losses.

Denuit [28] criticized Bühlmann's studies due to the mathematical complexity and defended the bonus-malus system for individual-based calculations. Bonus-malus systems basically generate premiums with individual risk factors in claim history, premiums can be affected by risk profile and past claim history.

In 2008, Wen et al [36] aims to extend the Bühlmann and Bühlmann-Straub credibility models that consider risk dependency with common effects. Gomez-Deniz [17] proposed Bühlmann's credibility premiums based on a weighted balanced loss function with different likelihoods and priors, and this approach aimed to generalize precedent models.

Payandeh and Najafabadi [26] suggested a functional solution by using the minimization of the mean squared error technique for modeling credibility for heavy-tailed distribution, which is the basic modeling problem in actuarial theory. In 2010, Linda and Kubanova [23] studied the application of the Bühlmann-Straub model for MTPL insurance in the Slovakian insurance industry.

Dornheim and Brazauskas [14] had embedded solutions that consist of mixed linear models, by this approach, when data fits log-location-scale distributions. Besides, Kim and Jeon [22] formed a credibility theory for truncated data and this study emphasized the sensitivity of credibility premium by varying the trimming threshold. Each defined threshold affects the credibility premium, which gives information about the relationship between deductible and pricing performance.

Pitselis [29] had theoretical innovations by using empirical linear Bayes estimation and derivation asymptotical optimality for credibility. In what follows, Pitselis [30] emphasized the relation between quantile values and credibility premiums and showed how affects the quantile values of the loss dataset to credibility premiums.

In 2016, Hassan Zadeh and Stanford [19] defined each contract as a risk parameter, and the phase type random variable is used for representing the time until absorption. They aimed to model both the initial probability vector and related transition matrix in the Markov chain approach, identify the risk parameter, and obtain a tractable likelihood function. Bühlmann credibility premiums are calculated by this modeling structure.

Based on the risk measure perspective of actuarial sciences, Pitselis [31] consolidated the credibility theory to risk measures calculations. Value at Risk (VaR), conditional tail expectation (CTE), tail conditional median, and quantile tail expectation are proposed risk measures in this study. As a result of this study, credible quantile tail expectation measures give robust results rather than other proposed risk measures.

Based on the previous study, Pitselis [32] extend this study by using a quantile regression framework, by this approach, the proposed model captures the risk of insurers individually. By this means, insurers that have the same riskiness have the same economical impact on portfolios and they can be grouped in the same class. In general, risk classification and risk diversification are significant factors for managing an insurance portfolio. So that managing the portfolio risk based on homogeneous risk classes can be more efficient.

Shi and Yang [33] embedded credibility theory with dependency structures based on pair copula constructions for insurance experience rating, They used mixed D-vine copula for dependency modeling of semi-continuous insurance claims. By this approach, they modeled the incorporation of past experience into future premiums.

Chukwudum [12] applied the empirical Bayesian credibility model for the estimation of risk premiums for various insurance branches in the Nigerian insurance industry. In 2019, for the Ghana insurance industry, Lotsi et al [24] studied the estimation of the Bühlmann-Straub frequency-severity claim model for non-life marine insurance. They suggested a credibility pricing model for companies that have insufficient claim history.

Diao and Weng [13] utilized machine learning techniques for credibility models and regression tree-based models were applied for the prediction of credibility premium with different covariates. Regression tree-based models have been used in actuarial sciences in pricing, risk classification, and reserving parts of the theory, however, this study was the first one for credibility modeling.

Winarta et al [38] proposed multivariate Bühlmann-Straub credibility on different lines of business for claim reserving, by this approach, the proposed model gives more accurate results than the standard credibility model. Yan and Song [40] extend Bühlmann credibility theory based on weighting credibility estimators with linear combinations optimally. In the proposed model, optimal weights have also asymptotically converging properties, these optimal weights gave remarkable and useful results based on standard finite-sample weights.

In Türkiye, credibility theory is used for both life and non-life insurance perspective. Selcuk-Kestel and Yıldırım Külekci [41] utilized Bühlmann's credibility for annuity net single premium modeling with consideration of longevity risk. They compared the German, Japanese, and Turkish mortality structures and credibility premiums at different ages for each country. Whereas, for non-life insurance, Bülbül and Baykal [11] used the credibility theory for the optimization of bonus-malus tariffs in the Turkish third-party liability insurance dataset. Obtained optimal bonus-malus rates are assumed as a negative binomial distribution.

CHAPTER 3

MTPL INSURANCE IN TÜRKİYE

In Türkiye, the number of compulsory insurance is 12, shown in Table 3.1. MTPL is one of the most significant insurances in Türkiye especially based on insurance volume [1]. Therefore, in MTPL insurance, positive or negative situations might directly affect the Türkiye economy.

Table 3.1: Compulsory insurances in Türkiye

Level	Compulsory Insurances List
1	Motor Third Party Liability Insurance (MTPL)
2	Compulsory Liability Insurance for Bottled Gas
3	Compulsory Personal Accident Insurance for Miners
4	Compulsory Financial Liability Insurance for Hazardous Materials and Hazardous Waste
5	Compulsory Liability Insurance for Marine Pollution by Waterside Facilities
6	Compulsory Liability Insurance for Medical Malpractice
7	Compulsory Seat-Based Personal Accident Insurance for Highways Passenger Transportation
8	Compulsory Financial Liability Insurance for Private Security
9	Compulsory Liability Insurance for Marine Vehicles
10	Compulsory Financial Liability Insurance for Certificate
11	Compulsory Foreign Travel Insurance
12	Compulsory Earthquake Insurance (TCIP)

The intention of MTPL insurance is to partly or entirely cover the loss caused to the car or the life of the "other party" in case of an accident, depending on the fault rate. In accordance with the Highways Law No. 2918 [4], it must be done by all motor vehicles on the highway. MTPL is a liability insurance and covers the material and bodily damages caused by vehicles on the road to third parties. It does not cover the damages that the vehicle drivers may cause to their own vehicle or themselves. It means that MTPL insurance does not replace motor own damage (MOD) insurance. MTPL insurance also does not cover moral claims. Another feature of MTPL insurance in Türkiye is that the policy is valid within the borders of Türkiye.

In accordance with the payment for MTPL insurance in Türkiye, the premium is paid in advance against the delivery of the policy as soon as the contract is concluded. However, the parties may agree to pay in installments, provided that at least one-quarter of the insurance premium is paid in advance in return for the delivery of the policy. In this case, the insurer is deemed to have waived its right to terminate the contract due to non-payment of premium.

3.1 General Coverages in MTPL Insurance

The types of MTPL coverages which are material damages coverage, permanent disability coverage, health expenses coverage, and compensation for loss of support can be summarized as:

i) Material damages coverage: It is the reduction in direct property, including the depreciation of the vehicle. Depreciation in vehicle accidents for which the insured is responsible is determined, upon request, by the insurance expert in the relevant branch.

ii) Permanent disability coverage: It is the guarantee to be determined to cover the financial losses that the third parties will suffer in the future due to their permanent disability. After the treatment of the injured is completed due to the accident, the caregiver expenses incurred after the permanent disability rate is determined by the disability medical board report to be obtained from an authorized hospital are within the scope of the permanent disability coverage, provided that they are limited to these coverage limits. In determining the amount of compensation in question, the disabled person is taken as a basis.

iii) Health expenses coverage: It is the guarantee that includes all treatment expenses, including the costs of prosthetic organs, in order to ensure that the third person is physically restored due to a traffic accident. From the beginning of the treatment of the victim due to the accident until the injured receives a permanent disability report, the caregiver's expenses incurred during the treatment, other expenses related to the treatment, and the expenses related to the partial or complete reduction of the working because of the traffic accident are within the scope of the health expense coverage. Health expenses coverage is the responsibility of the Social Security Institution of Türkiye.

iv) Compensation for loss of support: It is the compensation to be determined to cover the support losses of those who are deprived of the support of the deceased due to the death of the third person. The deceased person is taken as a basis for determining the amount of compensation in question. For example, one of the factors considered for compensation is the salary of the deceased.

3.2 Coverage Limits of MTPL Insurance

The upper limits of the coverages of MTPL are determined by the Undersecretariat of Treasury and these limits are kept constant. Therefore, it is not possible to change the limits arbitrarily. If the limits determined by the Undersecretariat of Treasury for compulsory MTPL are not sufficient to cover all the damage caused to the other party, the rest of the costs related to the damage will be paid by the policyholder.

Per-person health expenses and injury and death coverage limits in Türkiye are shown with their changes compared to the previous period in Figure 3.1 [2]. It is apparently seen that there is a huge increase from January 2022 to July 2022. In this period, coverage limits are doubled by the Undersecretariat of the Treasury. One of the most significant reasons for this is the high increase in inflation during these periods. If such a limit increase had not occurred, it is likely that the policyholders would have to pay most of the damage themselves, because the limits would have been exceeded usually.

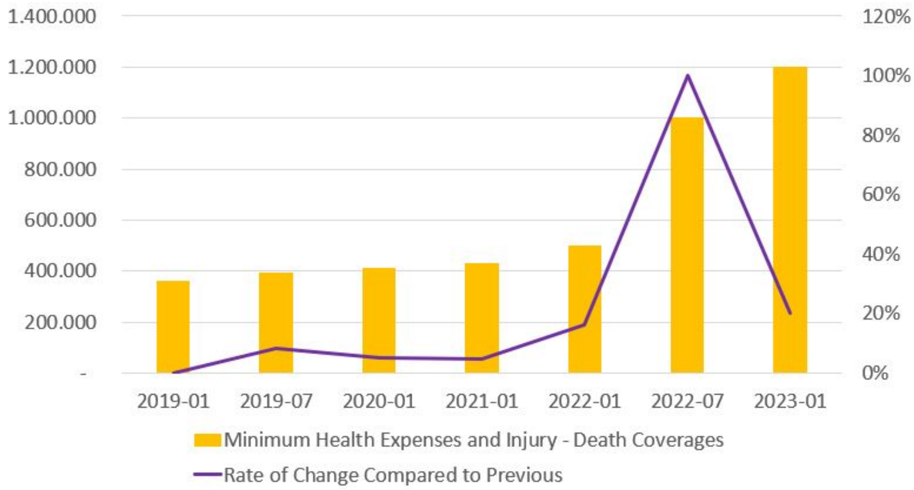

Figure 3.1: Minimum health expenses and injury - death coverages limit between 2019-2022

Coverage limits for material damages in MTPL are given per vehicle and per accident in Türkiye. Limits between 2019-2022 are illustrated in Figure 3.2 [2]. Same as in the health expenses and injury and death coverage limits, material damage coverage has shown a sharp increase in July 2022 due to high inflation.

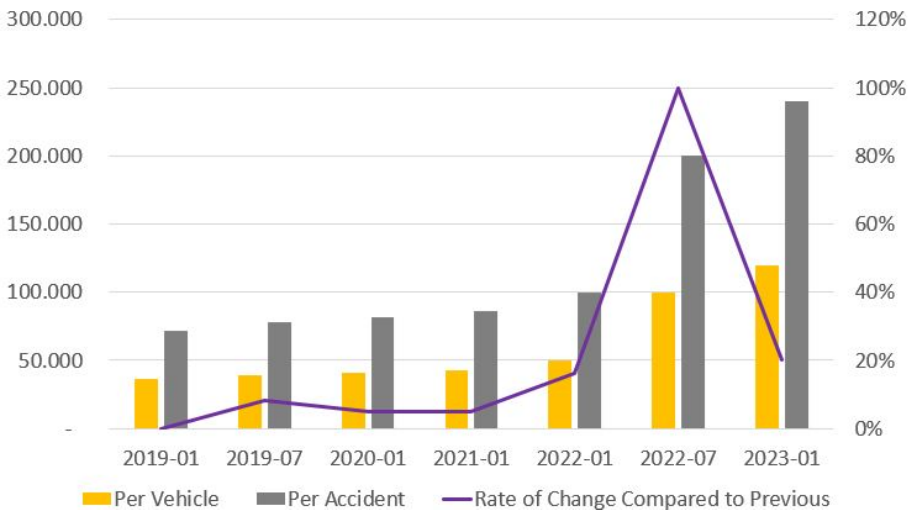

Figure 3.2: Material damage coverage limits between 2019-2022

3.3 Premium Increases in MTPL Insurance

There have been made increases in MTPL cap premiums in Türkiye announced by PPRSA [6]. These increases are made on a monthly basis. Table 3.2 shows the quarterly and cumulative changes in the cap values, from 2019 to 2022. The high increase seen in the coverage limits is also seen in the premium part. In the first half of 2022, the limits are increased by 100%. Considering the cap premium increases, high increases are seen in 2022. From the first quarter of 2019 to the end of 2022, the increase in the cap premiums is 155% in total.

Table 3.2: Quarterly premium increases

Year	Quarter	Quarterly Increase	Cumulative Increase
2019	Q1	4.57%	4.57%
	Q2	4.57%	9.35%
	Q3	4.57%	14.35%
	Q4	4.57%	19.57%
2020	Q1	4.57%	25.04%
	Q2	3.80%	29.79%
	Q3	2.27%	32.73%
	Q4	2.27%	35.75%
2021	Q1	3.03%	39.86%
	Q2	3.03%	44.10%
	Q3	3.03%	48.46%
	Q4	3.03%	52.96%
2022	Q1	24.86%	90.99%
	Q2	6.12%	102.68%
	Q3	9.52%	121.97%
	Q4	14.94%	155.13%

3.4 Cap Premium Structure and Türkiye Insurance Sector Results in MTPL Insurance

In Türkiye, cap premiums have been applied since 12 April 2017 for MTPL insurance [5]. After this date, the significance of calculating an accurate cap premium has also increased because it is compulsory regulation in Türkiye. The factors used in the computation of the cap premium can be listed as vehicle type, the city in which the vehicle is registered, and the NCD level of the vehicle owner.

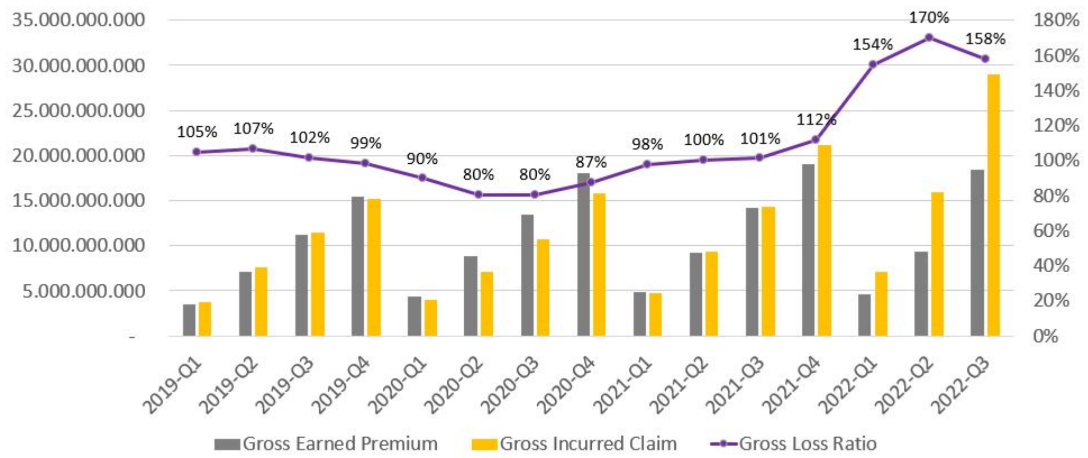

Figure 3.3: Gross loss ratios of MTPL insurance in Türkiye

It is clearly seen in Figure 3.3 that there are high differences between gross incurred claims, gross earned premiums, and gross loss ratios in comparison with years 2019 and 2022 [7]. Although gross earned premiums increased, gross incurred claims also increased and this increase is higher than gross earned premium. This situation ended up with higher gross loss ratios.

It is also seen that loss ratios are lower in the pandemic period than in other periods due to lower volumes of vehicle use. After the effects of the Covid-19 pandemic is alleviated, loss ratios increased sharply than previous periods with inflation effects as well. Also, the year 2022 results in Figure 3.3 apparently show the importance of cap premium estimation and pricing strategies because gross loss ratios are higher than 150% for the last three periods.

CHAPTER 4

CREDIBILITY METHODS

The classical approach has identified the significance of joining pieces of information from past experiences and the recent experiences of individuals. By combining data, better predictions can be accomplished instead of using just recent observations. For an example of this situation, when looking at the İstanbul values from the dataset used in this thesis, the average claim for the last quarter is around 1996 TL, moreover, the average claim for other quarters is also known. Therefore, instead of taking into account just recent observations, premiums for the next quarter can be found by using historical data in the light of credibility theory. To illustrate the example, Figure 4.1 is given below.

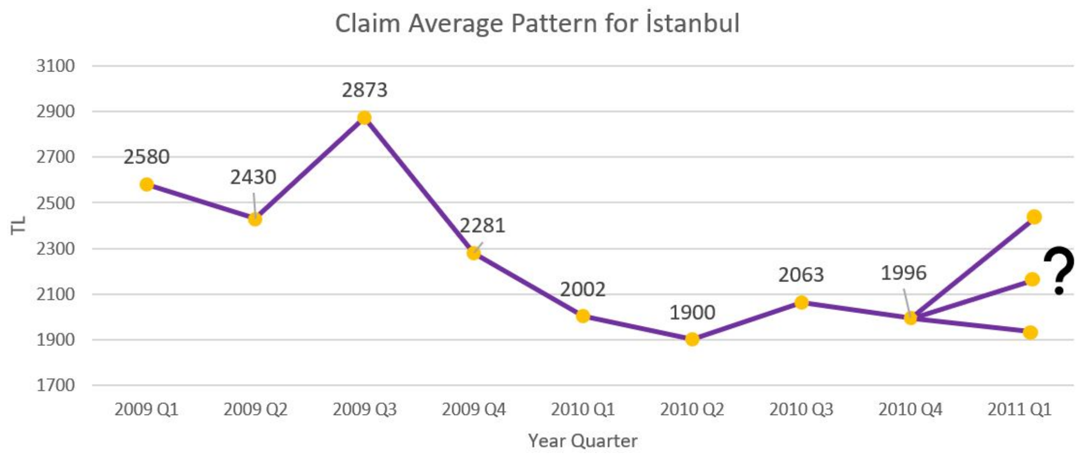

Figure 4.1: Illustration of credibility example

In this chapter, credibility theory is examined in three sub-titles which are classical credibility, Bühlmann credibility, and Bühlmann-Straub credibility methods. Classical credibility predicts the updated premium by using recent data and the manual rate.

In the Bühlmann credibility method is based on the least mean squared error (MSE) for optimal estimates, also, in this method variabilities between groups and in groups are identified. Bühlmann credibility is extended to Bühlmann-Straub credibility by using different level exposures to improve the forecast.

4.1 Classical Credibility

The classical approach known as limited-fluctuation credibility is the oldest approach. It suggests updating the loss estimation as a weighted average of the estimation based on recent data and the rate in the insurance manual.

In case, If the amount of recent data is adequate for the model, the data achieves full credibility, and the updated estimation can be based only on the recent data. However, if the amount of recent data is found to be inadequate, full credibility cannot be achieved, and partial credibility is attributed to the data. There are several claim experience measures that the classical credibility approach can be based on:

- Claim Severity
- Claim Frequency
- Aggregate Loss
- Pure Premium

To introduce the loss experience measures, each of them is defined separately as follows:

Claim Frequency: The number of claims denoted by N.

Claim Amount: It is denoted by X. In addition, each independent ith claim is represented by X_i.

Aggregate loss: It is expressed as a collection of claim amounts. It is denoted by S and $S = X_1 + X_2 + X_3 + \ldots + X_N$.

Claim Severity: The sample mean of claim amounts is the average claim severity. $X_1, X_2, \ldots, X_N$ i.e. $\bar{X} = S/N$.

Pure premium: It is denoted by P and let E be the number of exposure. $P = S/E$.

The updated prediction based on the credibility approach, U, is given as 4.1.

$$U = ZD + (1 - Z)M, \tag{4.1}$$

where,

M represents the historical data and estimation based on recent data represented by D. Z is the credibility factor that shows the credibility of recent data, where $0 \leq Z \leq 1$. U is the prediction for the next year.

4.1.1 Full Credibility

The classical credibility approach indicates the necessary minimum size of data for full credibility ($Z = 1$). This size is called the standard for full credibility. These standards are varied with claim experience measures.

4.1.1.1 Full Credibility for Claim Frequency

Under the normality assumption, the confidence intervals for claim frequency, N are given in 4.2. Based on this assumption N is distributed normally with μ and σ^2.

$$\begin{aligned} Pr(\mu_N - k\mu_N \leq N \leq \mu_N + k\mu_N) &= Pr\left(-\frac{k\mu_N}{\sigma_N} \leq \frac{N - \mu_N}{\sigma_N} \leq \frac{k\mu_N}{\sigma_N}\right) \\ &= \Phi\left(\frac{k\mu_N}{\sigma_N}\right) - \Phi\left(-\frac{k\mu_N}{\sigma_N}\right) \\ &= \Phi\left(\frac{k\mu_N}{\sigma_N}\right) - \left[1 - \Phi\left(\frac{k\mu_N}{\sigma_N}\right)\right] \\ &= 2\Phi\left(\frac{k\mu_N}{\sigma_N}\right) - 1. \end{aligned} \tag{4.2}$$

Then,

$$\frac{k\mu_N}{\sigma_N} = Z_{1-\frac{\alpha}{2}}. \tag{4.3}$$

The observed frequency is within $100k\%$ of the true mean with the probability of $(1-\alpha)$.

It is assumed that the claim frequency has a Poisson distribution which has enough large mean for a normal approximation to facilitate the calculations. Under the assumption, the Equation 4.3 is written as

$$\frac{k\lambda_N}{\sqrt{\lambda_N}} = k\sqrt{\lambda_N} = Z_{1-\frac{\alpha}{2}}. \tag{4.4}$$

Full credibility can be attained for frequency if $\lambda_N \geq \lambda_F$ where

$$\lambda_F \equiv \left(\frac{Z_{1-\frac{\alpha}{2}}}{k}\right)^2. \tag{4.5}$$

4.1.1.2 Full Credibility for Claim Severity

Suppose, there is N number of claim amounts that are independently and identically distributed (iid) with mean μ_X, variance σ_X^2, and $\bar{X}$ sample mean. Furthermore, N is enough large for the normal approximation. To attain the full credibility for $\bar{X}$.

$$\begin{aligned} Pr(\mu_X - k\mu_X \leq \bar{X} \leq \mu_X + k\mu_X) &= Pr\left(-\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}} \leq \frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{N}}} \leq \frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right) \\ &= \Phi\left(\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right) - \Phi\left(-\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right) \\ &= \Phi\left(\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right) - \left[1 - \Phi\left(\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right)\right] \\ &\simeq 2\Phi\left(\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right) - 1 \end{aligned} \tag{4.6}$$

so that,

$$\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}} \geq Z_{1-\frac{\alpha}{2}}, \tag{4.7}$$

and

$$\begin{aligned} N &\geq \left(\frac{Z_{1-\frac{\alpha}{2}}}{k}\right)^2 \left(\frac{\sigma_X}{\mu_X}\right)^2, \\ N &\geq \lambda_F C_X^2. \end{aligned} \tag{4.8}$$

where C_X^2 is the coefficient of variation.

In claim severity, the Poisson distribution assumption might not hold as in claim frequency. For implementation in practice, mean and variance must be calculated from the sample.

4.1.1.3 Full Credibility for Aggregate Loss

Until this part, full credibility requirements and calculations are given for individual claims. In this part, these are presented for aggregate loss.

The confidence intervals based on the normal approximation for aggregate loss, S are given in 4.9.

$$\begin{aligned} Pr(\mu_S - k\mu_S \leq S \leq \mu_S + k\mu_S) &= Pr\left(-\frac{k\mu_S}{\sigma_S} \leq \frac{S-\mu_S}{\sigma_S} \leq \frac{k\mu_S}{\sigma_S}\right) \\ &= \Phi\left(\frac{k\mu_S}{\sigma_S}\right) - \Phi\left(-\frac{k\mu_S}{\sigma_S}\right) \\ &= \Phi\left(\frac{k\mu_S}{\sigma_S}\right) - \left[1 - \Phi\left(\frac{k\mu_S}{\sigma_S}\right)\right] \\ &= 2\Phi\left(\frac{k\mu_S}{\sigma_S}\right) - 1. \end{aligned} \tag{4.9}$$

Let, N and X_i's be independent, and N has Poisson distribution to calculate the mean and variance of the compound distribution. To compute mean and variance,

$$\begin{aligned} E[S] =& E[N]E[X]. \\ \mu_S =& \mu_N \mu_X. \end{aligned} \tag{4.10}$$

$$\begin{aligned} Var(S) =& E[N]Var[X] + Var[N](E[X])^2. \\ \sigma_S^2 =& \mu_N \sigma_X^2 + \mu_X^2 \sigma_N^2. \end{aligned} \tag{4.11}$$

Thus,

$$\begin{aligned} \frac{\mu_S}{\sigma_S} =& \frac{\lambda_N \mu_X}{\sqrt{\lambda_N(\mu_X^2 + \sigma_X^2)}} \\ =& \frac{\mu_X \sqrt{\lambda_N}}{\sqrt{\mu_X^2 + \sigma_X^2}}. \end{aligned} \tag{4.12}$$

Therefore, Equation 4.9 is written as,

$$\begin{aligned} Pr(\mu_S - k\mu_S \leq S \leq \mu_S + k\mu_S) &= 2\Phi\left(\frac{k\mu_S}{\sigma_S}\right) - 1 \\ &= 2\Phi\left(\frac{k\mu_X\sqrt{\lambda_N}}{\sqrt{\mu_X^2 + \sigma_X^2}}\right) - 1. \end{aligned} \tag{4.13}$$

Probability should be at least $(1 - \alpha)$. So, it has to be

$$\frac{k\mu_X\sqrt{\lambda_N}}{\sqrt{\mu_X^2 + \sigma_X^2}} \geq Z_{1-\frac{\alpha}{2}}. \tag{4.14}$$

Then,

$$\lambda_N \geq \left(\frac{Z_{1-\frac{\alpha}{2}}}{k}\right)^2 \left(\frac{\mu_X^2 + \sigma_X^2}{\mu_X^2}\right). \tag{4.15}$$

The standard for full credibility

$$\left(\frac{Z_{1-\frac{\alpha}{2}}}{k}\right)^2 \left(\frac{\mu_X^2 + \sigma_X^2}{\mu_X^2}\right) = \lambda_F(1 + C_X^2), \tag{4.16}$$

$$\lambda_F(1 + C_X^2) = \lambda_F + \lambda_F C_X^2. \tag{4.17}$$

It is seen in Equation 4.17 the standard for aggregate loss is equal to the sum of the standards of claim frequency and claim severity.

4.1.1.4 Full Credibility for Pure Premium

Pure premium is equal to risk premium which is calculated without any loadings. The expected cost is computed by adding loadings to pure premium. After that, the technical premium is calculated by adding items such as profit, general expenses, and commission to this expected loss.

The confidence intervals based on the normal approximation for pure premium, P are given in 4.18.

$$\begin{aligned} Pr(\mu_P - k\mu_P \leq P \leq \mu_P + k\mu_P) &= Pr\left(-\frac{k\mu_P}{\sigma_P} \leq \frac{P - \mu_P}{\sigma_P} \leq \frac{k\mu_P}{\sigma_P}\right) \\ &= \Phi\left(\frac{k\mu_P}{\sigma_P}\right) - \Phi\left(-\frac{k\mu_P}{\sigma_P}\right) \\ &= \Phi\left(\frac{k\mu_P}{\sigma_P}\right) - \left[1 - \Phi\left(\frac{k\mu_P}{\sigma_P}\right)\right] \\ &= 2\Phi\left(\frac{k\mu_P}{\sigma_P}\right) - 1. \end{aligned} \tag{4.18}$$

Because of the fact that exposure E is constant, Equation 4.13 can be used to compute the probability in Equation 4.18. It is concluded that the full credibility standard is the same for pure premium and aggregate loss.

4.1.2 Partial Credibility

Full credibility cannot be achieved if there is no adequately large risk group. As a result, Z which is smaller than 1 has to be specified.

4.1.2.1 Partial Credibility for Claim Frequency

The confidence intervals for claim frequency, in normality assumption, are given for partial credibility in Equation 4.19.

$$\begin{aligned} Pr(Z\mu_N - k\mu_N \leq ZN \leq Z\mu_N + k\mu_N) &= Pr\left(-\frac{k\mu_N}{Z\sigma_N} \leq \frac{N-\mu_N}{Z\sigma_N} \leq \frac{k\mu_N}{Z\sigma_N}\right) \\ &= \Phi\left(\frac{k\mu_N}{Z\sigma_N}\right) - \Phi\left(-\frac{k\mu_N}{Z\sigma_N}\right) \\ &= \Phi\left(\frac{k\mu_N}{Z\sigma_N}\right) - \left[1 - \Phi\left(\frac{k\mu_N}{Z\sigma_N}\right)\right] \\ &= 2\Phi\left(\frac{k\mu_N}{Z\sigma_N}\right) - 1. \end{aligned} \tag{4.19}$$

Suppose claim frequency has Poisson distribution and the normal approximation is applied

$$2\Phi\left(\frac{k\mu_N}{Z\sigma_N}\right) - 1 = 2\Phi\left(\frac{k\sqrt{\lambda_N}}{Z}\right) - 1. \tag{4.20}$$

Thus,

$$\frac{k\sqrt{\lambda_N}}{Z} = Z_{1-\frac{\alpha}{2}}. \tag{4.21}$$

So,

$$\begin{aligned} Z &= \left(\frac{k}{Z_{1-\frac{\alpha}{2}}}\right) \\ &= \sqrt{\frac{\lambda_N}{\lambda_F}}. \end{aligned} \tag{4.22}$$

4.1.2.2 Partial Credibility for Claim Severity

With normality assumption, the confidence intervals for claim severity are given for partial credibility in Equation 4.23.

$$
\begin{aligned}
Pr(Z\mu_X - k\mu_X \leq Z\bar{X} \leq Z\mu_X + k\mu_X) &= Pr\left(-\frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}} \leq \frac{\bar{X}-\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}} \leq \frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}}\right) \\
&= \Phi\left(\frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}}\right) - \Phi\left(-\frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}}\right) \\
&= \Phi\left(\frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}}\right) - \left[1 - \Phi\left(\frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}}\right)\right] \\
&\simeq 2\Phi\left(\frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}}\right) - 1.
\end{aligned}
\tag{4.23}
$$

Thus,

$$
\frac{k\mu_X}{\frac{Z\sigma_X}{\sqrt{N}}} = Z_{1-\frac{\alpha}{2}}. \tag{4.24}
$$

Therefore,

$$
Z = \sqrt{\frac{N}{\lambda_F C_X^2}}. \tag{4.25}
$$

4.1.2.3 Partial Credibility for Aggregate Loss

The confidence interval for partial credibility in accordance with an aggregate loss is given in Equation 4.26.

$$
\begin{aligned}
Pr(Z\mu_S - k\mu_S \leq ZS \leq Z\mu_S + k\mu_S) &= Pr\left(-\frac{k\mu_S}{Z\sigma_S} \leq \frac{S-\mu_S}{Z\sigma_S} \leq \frac{k\mu_S}{Z\sigma_S}\right) \\
&= \Phi\left(\frac{k\mu_S}{Z\sigma_S}\right) - \Phi\left(-\frac{k\mu_S}{Z\sigma_S}\right) \\
&= \Phi\left(\frac{k\mu_S}{Z\sigma_S}\right) - \left[1 - \Phi\left(\frac{k\mu_S}{Z\sigma_S}\right)\right] \\
&= 2\Phi\left(\frac{k\mu_S}{Z\sigma_S}\right) - 1.
\end{aligned}
\tag{4.26}
$$

Thus,

$$\frac{k\mu_X\sqrt{\lambda_N}}{Z\sqrt{\mu_X^2+\sigma_X^2}} = Z_{1-\frac{\alpha}{2}}. \tag{4.27}$$

So, we get

$$Z = \sqrt{\frac{\lambda_N}{\lambda_F(1+C_X^2)}}. \tag{4.28}$$

4.1.2.4 Partial Credibility for Pure Premium

Under the normality assumption, The confidence interval for partial credibility in accordance with a pure premium is given in Equation 4.29.

$$\begin{aligned} Pr(Z\mu_P - k\mu_P \leq ZP \leq Z\mu_P + k\mu_P) &= Pr\left(-\frac{k\mu_P}{Z\sigma_P} \leq \frac{P-\mu_P}{Z\sigma_P} \leq \frac{k\mu_P}{Z\sigma_P}\right) \\ &= \Phi\left(\frac{k\mu_P}{Z\sigma_P}\right) - \Phi\left(-\frac{k\mu_P}{Z\sigma_P}\right) \\ &= \Phi\left(\frac{k\mu_P}{Z\sigma_P}\right) - \left[1-\Phi\left(\frac{k\mu_P}{Z\sigma_P}\right)\right] \\ &= 2\Phi\left(\frac{k\mu_P}{Z\sigma_P}\right) - 1. \end{aligned} \tag{4.29}$$

So, we get

$$Z = \sqrt{\frac{\lambda_N}{\lambda_F(1+C_X^2)}}. \tag{4.30}$$

4.2 Bühlmann Credibility

In Bühlmann models, the interaction between the variation of group distinctions and the variation of within-group fluctuations is identified. In this model, the exposure measure is assumed the equal for all groups.

The loss measure is denoted by X for insurance policies block or a risk group. X may be claim severity, aggregate loss, pure premium or claim frequency. It is assumed that the group risk profiles are identified by a parameter that is θ. θ defines the distribution of X. Conditional mean and variance can be denoted as

$$\mathrm{E}(X|\theta) = \mu_X(\theta). \tag{4.31}$$

$$\mathrm{Var}(X|\theta) = \sigma_X^2(\theta). \tag{4.32}$$

It is assumed that insurance companies have similar policy blocks with various risk profiles. In this way, groups are varied with the parameter of θ. The distribution of Θ is called the prior distribution. Consequently, the conditional mean and variance of X turn into random variables in Θ. They are denoted by $\mu_X(\Theta) = \mathrm{E}(X|\Theta)$ and $\sigma_X^2(\Theta) = \mathrm{Var}(X|\Theta)$, respectively.

In the Bühlmann model, it is also assumed that there are n observations and they are (iid) and depend on θ. Furthermore, n also can be a period of time. The aim is to predict of X for X_{n+1} based on X.

4.2.1 Variance Components

X has two variation components. One of them is between risk groups, and the second one is within risk groups.

The unconditional mean of X is

$$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X|\Theta)] = \mathrm{E}[\mu_X(\Theta)]. \tag{4.33}$$

The unconditional variance of X is

$$\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X|\Theta)] + \mathrm{Var}[\mathrm{E}(X|\Theta)]. \tag{4.34}$$

where $\mathrm{Var}(X|\Theta)$ is process variance and $\mathrm{E}(X|\Theta)$ is hypothetical mean. So, $\mathrm{E}[\mathrm{Var}(X|\Theta)]$ is called the expected value of the process variance, and $\mathrm{Var}[\mathrm{E}(X|\Theta)]$ is called the variance of the hypothetical means. Generally, hypothetical mean indicates the averages of claim frequency, aggregate claim, or claim severity of an individual combination of risk characteristics. The hypothetical refers to the conditional expectation, given that the specific risk characteristics combination. On the other hand, the process variance, as is seen from the formula, is the conditional variance, given the particular risk characteristics combination [20].

For the notation,

$$\text{E}[\text{Var}(X|\Theta)] = E[\sigma_X^2(\Theta)] = \mu_{\text{PV}}, \tag{4.35}$$

and

$$\text{Var}[\text{E}(X|\Theta)] = \text{Var}[\mu_X(\Theta)] = \sigma_{\text{HM}}^2. \tag{4.36}$$

So, Equation 4.34 turns into

$$\text{Var}(X) = \mu_{\text{PV}} + \sigma_{\text{HM}}^2. \tag{4.37}$$

For the empirical implementation of Bühlmann credibility, unbiased estimators of μ_{PV} and σ_{HM}^2 are given below.

$$\hat{\mu}_{\text{PV}} = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2}{\sum_{i=1}^{r}(n_i - 1)}, \tag{4.38}$$

where r denotes the number of risk groups, and X_{ij} represents the loss observation for $i = 1, \ldots, r$ and $j = 1, \ldots, n_i$.

$$\hat{\sigma}_{\text{HM}}^2 = \frac{\left[\sum_{i=1}^{r} n_i(\bar{X}_i - \bar{X})^2\right] - (r-1)\hat{\mu}_{\text{PV}}}{n - \frac{1}{n}\sum_{i=1}^{r} {n_i}^2}, \tag{4.39}$$

where $n = \sum_{i=1}^{r} n_i$.

If the sample sizes of groups are the same, the Equation 4.39 converts;

$$\hat{\sigma}_{\text{HM}}^2 = \frac{1}{r(n_i - 1)}\left(\sum_{i=1}^{r}(\bar{X}_i - \bar{X})^2\right) - \frac{\hat{\mu}_{\text{PV}}}{n_i} \tag{4.40}$$

Different parts of the total variance are measured with the expected value of the process variance and the variance of the hypothetical means. Homogeneous risk groups indicate that there are no big differences between the loss claims within the group. It means that conditional variance is small. If there are different risk profiles, it means that the variance of the hypothetical means will be large. If the above statements are seen in the portfolio, distinguishing the risk groups will be easier.

k is defined as

$$k = \frac{\mu_{\text{PV}}}{\sigma^2_{\text{HM}}}. \tag{4.41}$$

Bühlmann credibility model was developed by Bühlmann in 1967. The approach relies on the linear predictor which uses past experiences to update the prediction of loss measure. As in Classical Credibility, the updated prediction is calculated with the formula below [34].

$$U = ZD + (1 - Z)M. \tag{4.42}$$

where D is the sample mean that comes from the data and M is the overall prior mean.

Bühlmann credibility factor depends on the number of observations and the ratio of EPV to VHM which is equal to k. Sample size and the credibility factor are directly proportional which means when the number of observations increases, the credibility factor also approximates 1. It is shown in Figure 4.2.

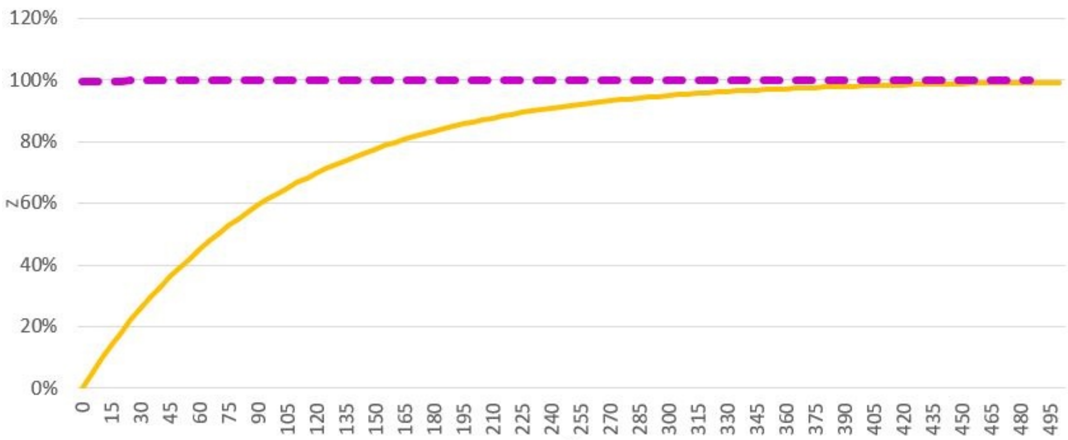

Figure 4.2: Credibility factor and the number of claims are directly proportional.

The first assumption of the Bühlmann model is that X's are identically and independently distributed. Furthermore, this distribution depends on the parameter θ [34].

Secondly, the parameter of θ depends on the random variable of Θ. The conditional variance and mean of random variable X as below [9].

$$\mathrm{E}(X|\theta) = \mu_X(\theta), \tag{4.43}$$

and

$$\mathrm{Var}(X|\theta) = \sigma_X^2(\theta). \tag{4.44}$$

The main aim of Bühlmann credibility is to predict X_{n+1} which relies on the linear function of X. There is an assumption that the distribution of X_{n+1} is the same as the distribution of X.

In the prediction of X_{n+1}, the used predictor which is $\hat{X}_{n+1}$ minimizes the mean squared error regarding the joint distribution of X_{n+1}, X and Θ. The predictor is

$$\hat{X}_{n+1} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n. \tag{4.45}$$

In Equation 4.45, all β coefficients are chosen to minimize the mean squared error which is formulated as

$$\mathrm{MSE} = \mathrm{E}[(X_{n+1} - \hat{X}_{n+1})^2]. \tag{4.46}$$

As a result, the Bühlmann premium is calculated as

$$\begin{aligned} \hat{X}_{n+1} &= \hat{\beta}_0 + \hat{\beta}'_S X, \\ \hat{X}_{n+1} &= \frac{n\bar{X}}{n+k} + \frac{k\mu_X}{n+k}, \\ \hat{X}_{n+1} &= Z\bar{X} + (1-Z)\mu_X \end{aligned} \tag{4.47}$$

where,

$$Z = \frac{n}{n+k}, \tag{4.48}$$

and

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{4.49}$$

In Equation 4.48, Z is called the Bühlmann credibility factor, and k is called the parameter of Bühlmann credibility.

4.3 Bühlmann–Straub Credibility

The Bühlmann-Straub model was developed by Bühlmann and Straub in 1970, firstly, to specify the claim ratios for the reinsurance market. After that, this model was widely used for other issues in the insurance sector [9]. As an example of usage areas, pricing calculation and classification to find true pure premiums for health and motor insurance can be given. In the Bühlmann credibility, loss observations are assumed as identically distributed with the same variance [21]. This assumption is pointless when the data comes from different periods and loss observations are not identically distributed.

In the Bühlmann-Straub model, the loss measure process variance is supposed to depend on the exposure. The exposure is denoted by m_i and the loss which comes from the different units of exposure is X_i. It is also significant that the exposure does not need to be only the number of insured. The unit of exposure can be the number of months, the number of quarters, and the received amount of premiums in the year i. As mentioned before, the conditional variance of loss distributions alters with the unit of exposure m_i, such that

$$\text{Var}(X_i|\Theta) = \frac{\sigma_X^2(\Theta)}{m_i}. \tag{4.50}$$

The first assumption of the Bühlmann-Straub model is that loss variable X is distributed not identically but independently. Also, this distribution depends on θ. Secondly, the parameter θ is also a random variable of Θ. Therefore, the mean and variance of X are defined as the conditional

$$\text{E}(X_i|\theta) = \mu_X(\theta), \tag{4.51}$$

and

$$\text{Var}(X_i|\theta) = \frac{\sigma_X^2(\theta)}{m_i}. \tag{4.52}$$

The unconditional mean is

$$\text{E}(X_i) = \text{E}[\text{E}(X_i|\Theta)] = \text{E}[\mu_X(\Theta)] = \mu_X. \tag{4.53}$$

The conditional variance mean is

$$\begin{aligned}
\mathrm{E}[\mathrm{Var}(X_i|\Theta)] &= \mathrm{E}\left[\frac{\sigma_X^2(\Theta)}{m_i}\right], \\
\mathrm{E}[\mathrm{Var}(X_i|\Theta)] &= \frac{\mu_{\text{PV}}}{m_i}.
\end{aligned} \tag{4.54}$$

The variance of the conditional mean is

$$\begin{aligned}
\mathrm{Var}[\mathrm{E}(X_i|\Theta)] &= \mathrm{Var}[\mu_X(\Theta)], \\
\mathrm{Var}[\mathrm{E}(X_i|\Theta)] &= \sigma_{\text{HM}}^2.
\end{aligned} \tag{4.55}$$

So, the total variance can be written as

$$\begin{aligned}
\mathrm{Var}(X_i) &= \mathrm{E}[\mathrm{Var}(X_i|\Theta)] + \mathrm{Var},[E(X_i|\Theta)], \\
\mathrm{Var}(X_i) &= \frac{\mu_{\text{PV}}}{m_i} + \sigma_{\text{HM}}^2.
\end{aligned} \tag{4.56}$$

For the empirical implementation of Bühlmann-Straub credibility, unbiased estimators of μ_{PV} and σ_{HM}^2 are given below.

$$\hat{\mu}_{\text{PV}} = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij}(X_{ij} - \bar{X}_i)^2}{\sum_{i=1}^{r}(n_i - 1)}, \tag{4.57}$$

and

$$\hat{\sigma}_{\text{HM}}^2 = \frac{\left[\sum_{i=1}^{r} m_i(\bar{X}_i - \bar{X})^2\right] - (r-1)\hat{\mu}_{\text{PV}}}{m - \frac{1}{m}\sum_{i=1}^{r} {m_i}^2}. \tag{4.58}$$

The predictor of Bühlmann-Straub minimizes the mean squared error of X_{n+1} predictors which are linear in X regarding the joint distribution of Θ, X_{n+1}, and X, as same as the Bühlmann model. The estimator of X_{n+1}, $\hat{X}_{n+1}$ is given as

$$\hat{X}_{n+1} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n. \tag{4.59}$$

In Equation 4.59, all β coefficients are chosen to minimize the mean squared error.

The formulation of the Bühlmann-Straub premium is the same as the Bühlmann premium except for the variance term.

$$\begin{aligned}
\hat{X}_{n+1} &= \hat{\beta}_0 + \hat{\beta}_S^{'} X, \\
\hat{X}_{n+1} &= Z\bar{X} + (1 - Z)\mu_X,
\end{aligned} \tag{4.60}$$

where

$$Z = \frac{m}{m+k}, \tag{4.61}$$

and

$$m = \sum_{i=1}^{n} m_i, \tag{4.62}$$

and

$$\bar{X} = \frac{1}{m} \sum_{i=1}^{n} m_i X_i. \tag{4.63}$$

If all exposures are the same for all periods, Equation 4.63 can be written as

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i. \tag{4.64}$$

CHAPTER 5

K-MEANS AND HIERARCHICAL CLUSTERING ALGORITHMS

Clustering is a wide term that refers to a variety of strategies for identifying subgroups, or clusters, in a data collection. When the observations of a dataset are clustered, it is sought to partition them into distinct groups so that the observations within each group are quite similar to each other, while observations in different groups are quite different from each other.

Clustering is characterized as an unsupervised learning approach since there is no input data to compare the clustering algorithm's output to the true labels for evaluation of its performance.

5.1 K-Means Clustering Algorithm

The k-means algorithm is one of the most used and most common of these unsupervised learning methods for clustering. The k-means algorithm is an iterative algorithm and the underlying purpose is to assign each data to only one group and to separate the data into non-intersecting groups [39].

While the k-means algorithm divides the data into clusters, it creates clusters in such a way that the sum of the squared distance between the centroid of the cluster and the data points is minimal. The smaller the difference in the clusters, the more homogeneous clusters are obtained.

Let C denote sets having the indices of the data points in each group. The following two properties are satisfied by the sets:

In Equation 5.1, it is expressed that each data point is belonging to at least one cluster.

$$C_1 \cup C_2 \cup C_3 \cup \ldots \cup C_k = \{1, 2, 3, \ldots, n\}. \tag{5.1}$$

In Equation 5.2, it can be seen that any observation cannot be assigned to more than one cluster.

$$C_k \cap C_{k'} = \emptyset \qquad \text{for all } k \neq k^{'}. \tag{5.2}$$

$W(C_k)$ is the measurement of the cluster variation for C_k. When $W(C_k)$ is decreased, the homogeneity within groups is increased.

Equation 5.3 expresses that it is seek to minimize the variation within-cluster while partitioning the observations into k clusters.

$$\underset{C_1,\ldots,C_K}{\text{minimize}} \left\{ \sum_{k=1}^{K} W(C_k) \right\}. \tag{5.3}$$

Within-cluster variation should be defined to solve Equation 5.3. For this concept, squared Euclidean distance is the most common choice [16]. We define

$$W(C_k) = \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2, \tag{5.4}$$

where $|C_k|$ is the number of observations in the kth cluster.

Variation within-cluster for the kth cluster can be described as the sum of all of the pairwise squared Euclidean distances for each data point in the cluster, divided by the sum of the number of observations in the same cluster.

Combining Equations 5.3 and 5.4 gives

$$\underset{C_1,\ldots,C_K}{\text{minimize}} \left\{ \sum_{k=1}^{K} \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 \right\}. \tag{5.5}$$

5.1.1 Methods for Determining the Number of Clusters

In this part, it will be explained how and by which methods the number of k clusters can be determined for the k-means clustering algorithm. As a method, Elbow and Silhouette metrics will be discussed.

5.1.1.1 Elbow Method

The elbow method indicates a concept of what k might be based on the sum of the squared distance between the center point of the assigned clusters and the data points. As the name suggests, the point at which the sum of squared distance shows an elbow can be selected as the number of clusters k.

5.1.1.2 Silhouette Method

The silhouette score is the measure of how observation is similar within-cluster in comparison with other clusters. The silhouette coefficient for ith data point is denoted by S_i such that,

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)}, \tag{5.6}$$

where a_i denotes the average distance between the data point i and all the observations in the cluster i belongs. b_i is the average distance between observation i and all clusters which are different from the cluster of observation i.

$[-1, 1]$ is the domain of the silhouette coefficient.

i) If S_i = -1, the sample is assigned to the wrong clusters.

ii) If S_i = 0, the neighboring clusters are very close to the sample.

iii) If S_i = 1, the neighboring clusters are far away from the sample.

Therefore, it can be said that to have correct and well-separated clusters, the silhouette coefficient should be as high as possible.

5.2 Hierarchical Clustering Algorithm

The main distinction between k-means and hierarchical clustering methods is that pre-specified k is not needed in hierarchical clustering [16]. Hierarchical clustering can be applied in different ways by agglomerative and divisive hierarchical clustering methods. The principle of the agglomerative method is that similar clusters are sequentially merged. In comparison, the divisive method is applied by splitting observations grouped in a single cluster into clusters recursively. In this thesis, we apply the agglomerative clustering method.

In hierarchical clustering, the distance measure is used for determining the similarity of clusters as used in k-means clustering. The most common measure is the Euclidean distance which is also used in k-means clustering. In this thesis, the similarity is measured by the Euclidean distance for both clustering methods.

In hierarchical clustering, it is requisite to identify the start and end points of the distance measurement and it is determined by linkage criteria, such as single linkage, complete linkage, average linkage, and so on. Complete and average linkage are preferred because they give more balanced dendrograms than others generally [16]. In this thesis, a complete linkage that calculates pairwise dissimilarities between the observations in different clusters and takes the largest of those is used.

To determine the clusters and examine the relationship between observations, dendrograms are commonly used in hierarchical clustering. Moreover, the dendrogram can be thought of as the distance matrix summary. Therefore, It is significant to know that the dendrogram provides general information and foresight in terms of evaluation, not a definitive clustering result.

5.3 Advantages and Disadvantages of K-means and Hierarchical Clustering Methods

For the implementation both are straightforward, however, the k-means algorithm may be faster than hierarchical clustering computationally, especially for small k values. In both methods, the significant drawback is results are impacted by initial seeds.

The estimation of k which is pre-defined is more difficult in the k-means algorithm. However, the output of hierarchical clustering is more informative for k prediction thanks to the dendrogram. Moreover, the k-means algorithm is sensitive to scale which means that results might be changed as a result of scaling.

CHAPTER 6

EMPIRICAL ANALYSIS

This section presents the application of Bühlmann and Bühlmann-Straub methods in the MTPL dataset of Türkiye which is provided by the Insurance Association of Türkiye (IAT). for the city-basis premium calculation. Premiums are calculated accident year-quarter basis from 2009 to 2010. Results are compared with the first quarter of 2011 (2011-Q1). In calculations, cities are divided into groups by different methods.

The first perspective of grouping is that cities are ordered in accordance with the number of claims and divided into groups. In this analysis, the minimum and the maximum number of groups are 3 and 10, respectively. The motivation for grouping by the number of claims is to create homogenous groups in accordance with the number of claims.

In second grouping is created by the geographical regions of Türkiye. In this grouping, the number of claims or claim average has not been taken into account.

Moreover, in the last grouping methods, the k-means and hierarchical clustering algorithms are used to divide cities into groups. The motivation for using these algorithms which use numeric input for clustering is that groups are created by scaled claim averages and claim numbers. In k-means the number of clusters k is selected according to the results of elbow and silhouette methods. The number of clusters is selected in the light of the dendrogram for hierarchical clustering. As a result of these selections cities are divided into 3, 6, and, 9 groups by k-means and from 2 to 7 groups by hierarchical clustering methods.

Premium calculations with Bühlmann and Bühlmann-Straub credibility are done in R [15]. The general information of data is described below.

6.1 Data Description

Before starting the analysis and modeling, the datasets from the underwriting (UW) years 2009 and 2010 are merged into a single dataset. The raw data consists of 7,415,007 rows which represent the number of policies after combining the datasets. Analyzed raw data columns before generating new columns are shown in Table 6.1.

Table 6.1: Analyzed data columns

Column Names	Description of Columns
Policy No	Unique for policies
Policy Inception Date	Start date of the policy (UW basis)
Policy End Date	End date of the policy (UW basis)
Plate Code	Cities generated by plate codes
Model Year	Model year for vehicles
Capacity	Capacity of vehicles
Insured Nationality	Nationality of insureds
Claim Amount 1 Claim Amount 2 Claim Amount 3 Claim Amount 4 Claim Amount 5	Used for total claim amount
Total Claim Amount	Sum of claim amounts in Turkish Lira (TL)

Table 6.2 illustrates the new columns which are added for doing a more detailed analysis.

Table 6.2: Generated columns from the raw data

Column Names	Description of Columns
Capacity Class	Grouped capacities
City	Generated from plate codes
Claim Flag	Generated for the distinction of claimed policies
Inflation Loading Factor	Added for applying inflations on claim amounts
Claim Year-Quarter	Added for pricing based on the accident year-quarter
Region	Added for grouping based on the geographical region
Total Claim Amount Inf	Total claim amount adjusted for inflation
Total Claim Amount Inf Class	Added for distribution of inflated total claim amount

A capacity class is generated to check the distribution of policy count for capacity. The city column was also created by using the plate code column to see the credibility results in accordance with cities. Moreover, while creating a city column from plate codes, it is assumed that vehicles are used where their plate codes belong. The claim flag column is created by using the total claim amount column. If the total claim amount is greater than zero, the claim flag is assigned as one if not, zero. Also, the claim flag column is used to calculate the claim count. The inflation loading factor column is used for the inflation adjustment. By using these factors, the impact of inflation is removed from all total claim amounts. Inflation is applied to the data by the claim date. Applied inflation factors are indicated in Table A.1 [3]. The main idea of generating the claim year-month column is to apply the inflation loading factors to the total claim amounts row by row. Thanks to the claim year-month column, inflation factors from the other dataset are joined to the dataset. Furthermore, the claim year-quarter column is generated to use in the Bühlmann premium calculation. We include a region column to check the distribution of the cities in the data as well as clustering the cities for premium calculation. The total claim amount inf column is generated as the inflation loading factor multiplied by the total claim amount. The total claim amount inf class column is created to analyze the distribution of claim amounts.

Before making any filtration the number of observations and the percentage of occurrence in accordance with the inception year-quarter is illustrated below (UW basis).

As can be seen in Table 6.3, the maximum number of policies was produced in the second quarter of the year 2010. In contrast, the minimum number of policies was written in the first quarter of the year 2009.

Table 6.3: The policy count by inception year-quarter

Inception Year	Quarter	Policy Count	Percentage of Policy Count
2009	Q1	794,004	10.7%
	Q2	1,013,516	13.7%
	Q3	876,479	11.8%
	Q4	932,704	12.6%
2010	Q1	839,814	11.3%
	Q2	1,038,333	14.0%
	Q3	929,861	12.5%
	Q4	990,296	13.4%
Total		7,415,007	100.0%

6.2 Data Filtering and Cleaning

The last step before starting to analyze is the data filtration part. Data columns have some missing and misregistered values which might impact the result negatively.

The reason for the first filtration is that the dataset contains some claim dates which are earlier than the policy inception date. As this is not possible, 42 data rows having this situation are excluded from the dataset.

The second filtration is made for plate code. The dataset includes plate codes from 1 to 81. Apart from those, there is also the plate code 999 which does not match any city. These rows are not used in the analysis. The number of rows with a 999 plate code is 4410 in the dataset.

The third filtration is about the model year. The filtration is done in the models for greater than the year 2010 and smaller than the year 1990. The main reason for this filtration is that in the dataset, some vehicles have a model year such as 1900, 1910, and so on. Therefore,1990 is assigned as the floor for the model year. As a result of filtration, the number of rows which is in line with the situation is 6243. Therefore, those are excluded.

The filtration of data is continued with the capacity. Capacity filtration is made for increasing the accuracy of the dataset because there are rows that have a capacity of more than 60. Thus, 146 rows are excluded.

The last filtration is done with the total claim amount, which is thought to be a wrong entry. In the dataset, there is a claim which is the amount of 99,000,500 which is not possible to occur. It is excluded before performing the credibility modeling.

Consequently, the number of rows has decreased by 351,304 and declined to 7,063,703. After the filtration steps, there are no missing values in any column in the dataset.

In Table 6.4, the policy count and the percentage by policy inception year-quarter are illustrated. It can be seen in Table 6.3, the maximum policy is written in 2010-Q2. In comparison, the minimum number of policies is written in 2019-Q1. These indicators are in line with the number of policies before filtration. However, the percentage of occurrences has changed. Before filtration, the percentage of written policies in 2010-Q2 is 14% and it has increased to 14.4% with filtering. In 2019-Q1, the ratio of written policies is 10.7% and has risen to 11.2% after filtering.

Table 6.4: The policy count after filtration steps by inception year-quarter

Inception Year	Quarter	Policy Count	Percentage of Policy Count
2009	Q1	789,343	11.2%
	Q2	1,008,977	14.3%
	Q3	872,119	12.3%
	Q4	924,998	13.1%
2010	Q1	828,926	11.7%
	Q2	1,019,879	14.4%
	Q3	825,364	11.7%
	Q4	794,097	11.2%
Total		7,063,703	100.0%

6.3 Descriptive Statistics for MTPL Dataset

After all the filtering steps, in this part, descriptive statistics are presented. These statistics are shown on the underwriting year basis and the accident year basis. Credibility models are applied on an accident-year basis.

6.3.1 Descriptive Statistics in Accordance with Underwriting Year

In general, descriptive statistics for categorical variables are shown in terms of claim frequency, claim number, claim average, and policy count. Firstly, descriptive statistics on an underwriting year basis are illustrated.

The percentage of capacity class 0-5 in the MTPL dataset accounted for 98.3% of capacity class. Statistical indicators are shown in Table 6.5 for capacity class.

Table 6.5: Statistics for capacity

Model Year	Policy Count	Claim Count	Claim Frequency	Claim Severity Average	Claim Severity Std. Dev.
0-5	6,945,529	1825379	26.28%	2,015	4,929
6-10	22,927	6623	28.89%	1,726	3,697
11-15	49,901	14,280	28.62%	1,509	3,342
16-20	15,975	4,914	30.76%	1,687	3,454
21-35	19,934	5,606	28.12%	1,610	3,404
35+	9,437	3,301	34.98%	3,044	8,426
Total	7,063,703	1,860,103	26.33%	2,010	4,916

In the dataset, the model years are between 1990 and 2010. The maximum number of policy counts belongs to 2006 and its claim frequency is accounted for 26.55%. In terms of frequency, the maximum is 28.28% and it comes from the 2010 model year. These are shown in Table 6.6

Table 6.6: Statistics for model year

Model Year	Policy Count	Claim Count	Claim Frequency	Claim Severity Average	Claim Severity Std. Dev.
1990-2000	724,013	157,187	21.71%	1,727	3,306
2001	296,380	70,026	23.63%	1,811	3,724
2002	103,727	25,432	24.52%	1,918	4,377
2003	210,181	54,374	25.87%	2,001	4,475
2004	778,779	204,446	26.25%	1,878	4,358
2005	824,610	215,300	26.11%	1,924	4,389
2006	1,036,966	275,359	26.55%	2,045	4,858
2007	748,575	199,552	26.66%	2,095	5,284
2008	913,181	257,638	28.21%	2,145	5,558
2009	788,014	219,974	27.91%	2,138	5,768
2010	639,277	180,815	28.28%	2,107	54,77
Total	7,063,703	1,860,103	26.33%	2,010	4,916

In terms of Nationality, the majority is Türkiye which is accounted for 99.84%. The statistics for Nationality are given below in Table 6.7.

Table 6.7: Statistics for nationality

Nationality	Policy Count	Claim Count	Claim Frequency	Claim Severity Average	Claim Severity Std. Dev.
Türkiye	7,052,085	1,857,945	26.35%	2,047	5,232
Others	11,618	2,158	18.57%	2,010	4,915
Total	7,063,703	1,860,103	26.33%	2,010	4,916

In the region column, the highest policy count is written in the Marmara Region, accounting for 43.50%. The reason for this situation is that İstanbul is in this region. The Central Anatolia Region follows the Marmara Region in these statistics, accounting for around 18.80%. Although the policy count percentage of The Southeastern Anatolia Region is 3.07%, the claim frequency is 32.79% which is the highest. Statistics are shown in Table 6.8.

Table 6.8: Statistics by region

Region	Policy Count	Claim Count	Claim Frequency	Claim Severity Average	Claim Severity Std. Dev.
Marmara	3,072,783	864,793	28.14%	2,093	4,999
Central Anatolia	1,327,274	325,536	24.53%	2,000	4,778
Aegean	973,626	253,552	26.04%	1,771	4,150
Mediterranean	783,527	184,199	23.51%	1,962	5,037
Black Sea	508,888	114,230	22.45%	2,111	5,613
Southeastern Anatolia	216,991	71,144	32.79%	1,813	5,225
Eastern Anatolia	180,614	46,649	25.83%	2,083	5,257
Total	7,063,703	1,860,103	26.33%	2,010	4,916

To analyze the inflated total claim amount, descriptive statistics are given in two different ways. First, we examine the whole dataset with claims and no-claim policies together. This will make a difference in the average claim, and the average claim to be assigned will be calculated for all policies in the data. In the second case, no-claim policies will be deducted from the dataset, and statistics will be obtained only on policies with claims. The histogram of the total claim amount is shown in Figure 6.1.

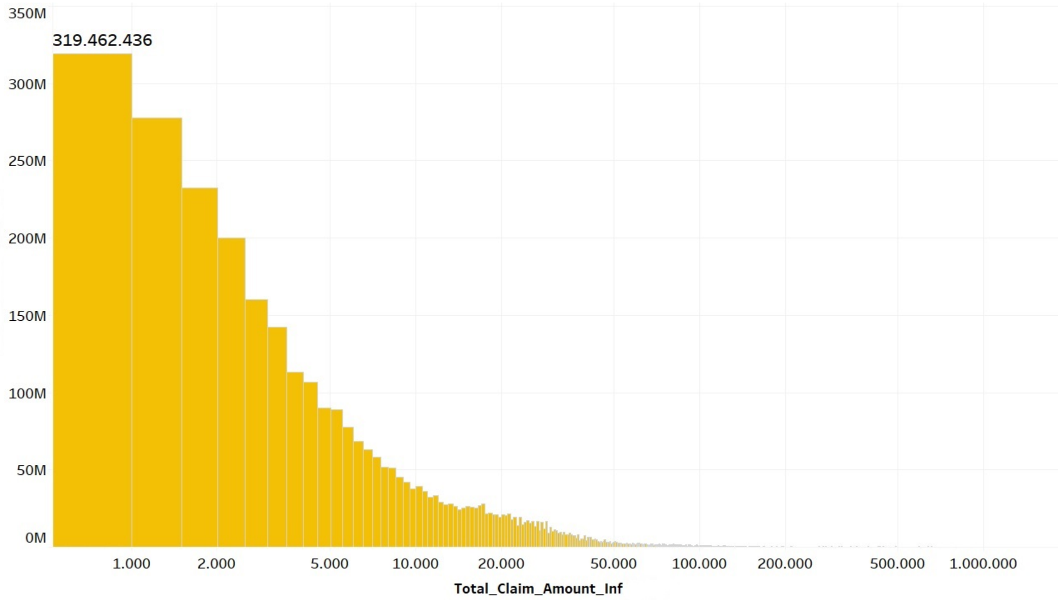

Figure 6.1: Histogram of inflated total claim amount

Table 6.9 illustrates that descriptive statistics of the whole data compared to the zero-claims excluded data changes, as expected. The average claim of policies with the claim is 1481 TL higher than the average of the whole data. Moreover, as well as average claims, other descriptive statistics such as 1st quarter, median, and 3rd quarter are higher for policies with claims.

Table 6.9: Descriptive statistics of inflated total claim amount

Descriptive Statistics	All Policies	Policies With Claim
1st Quarter	0	369
Median	0	779
Mean	529	2,010
3rd Quarter	182	1,761
Max	649,442	649,442
Standard Deviation	2,749	5,072
Coefficient of Variation	5.1965	2.5233

Descriptive statistics for cities are given in Table 6.10 and Table 6.11. Firstly, The dataset can be attributed as imbalanced due to the number of policies of İstanbul which accounted for 32.41% in the dataset. Moreover, the 5 cities with the highest number of policies are İstanbul, Ankara, İzmir, Bursa, and Antalya, respectively. This ranking is the same for the number of claims as expected. The 5 cities with the lowest number of policies are Tunceli, Hakkari, Kilis, Bayburt, and Gümüşhane, respectively. This order for these cities changes to Tunceli, Hakkari, Bayburt, Kilis, and Ardahan, respectively, for the number of claims.

The minimum claim frequency is 15.54% which comes from Ardahan. In comparison, Batman has the maximum claim frequency which is 42.76%. The dataset also shows that the highest 10 claim frequencies come from The Southeastern Anatolia Region and Eastern Anatolia Region.

In terms of claim severity average, the maximum and the minimum are 3231, and 1105 come from Hakkari and Kilis, respectively. When we check the lowest 10 claim severity averages, it is seen that 5 of them belong to Aegean Region. In comparison, 4 cities come from Eastern Anatolia Region out of the highest 10 cities.

These indicators show us that the claim severity average and claim frequency tend to increase as you move from the west to the east of Türkiye. These are also shown on Türkiye map in Figure A.1, and Figure A.3.

Table 6.10: Descriptive statistics for cities - 1

Region	City	Policy Count	Claim Count	Claim Frequency	Claim Severity Average	Claim Severity Std. Dev.
Aegean R.	Afyonkarahisar	32,705	7,769	23.85%	1,943	4,733
	Aydin	79,155	19,016	24.49%	1,668	3,948
	Denizli	84,334	21,971	26.51%	1,774	4,303
	İzmir	543,600	152,963	28.89%	1,746	3,933
	Kütahya	25,391	5,191	20.96%	2,123	5,387
	Manisa	82,917	19,351	23.23%	1,700	4,564
	Muğla	101,782	21,692	21.71%	1,966	4,704
	Uşak	23,742	5,599	23.65%	1,709	3,835
Black Sea R.	Amasya	22,069	4,434	20.26%	2,045	5,389
	Artvin	10,293	2,146	20.27%	2,576	7,049
	Bartın	8,062	1,604	19.75%	2,806	9,495
	Bayburt	2,504	528	21.07%	2,131	4,443
	Bolu	20,063	3,301	16.42%	2,294	5,952
	Çorum	32,693	7,325	22.01%	2,004	4,504
	Düzce	22,745	5,091	22.15%	2,354	4,630
	Giresun	28,050	6,655	23.84%	1,795	4,214
	Gümüşhane	4,467	902	20.88%	2,478	5,565
	Karabük	13,225	3,029	22.97%	1,779	4,885
	Kastamonu	19,672	4,507	22.38%	1,845	4,458
	Ordu	44,130	9,471	21.87%	2,185	5,196
	Rize	20,338	4,729	23.16%	2,388	8,418
	Samsun	96,084	22,685	23.34%	2,248	6,873
	Sinop	10,740	2,107	19.07%	2,361	5,423
	Tokat	31,342	5,883	18.06%	2,054	4,981
	Trabzon	73,338	16,582	22.66%	2,086	4,421
	Zonguldak	49,073	13,251	27.99%	1,819	4,988
Central Anatolia R.	Aksaray	16,594	3,451	20.45%	2,209	5,330
	Ankara	940,067	240,297	25.67%	1,987	4,702
	Çankırı	8,028	1,352	16.97%	1,825	4,131
	Eskişehir	73,756	17,814	24.79%	1,847	4,520
	Karaman	7,828	1,653	21.89%	1,839	4,750
	Kayseri	105,322	24,407	23.97%	2,017	4,676
	Kırıkkale	6,198	1,220	19.83%	2,175	5,179
	Kırşehir	10,355	2,003	19.64%	1,910	4,745
	Konya	79,241	17,341	21.35%	2,169	4,860
	Nevşehir	17,633	3,452	19.77%	2,091	4,957
	Niğde	16,589	3,445	20.93%	2,161	6,232
	Sivas	32,117	6,596	20.26%	1,958	4,758
	Yozgat	13,546	2,505	18.57%	2,622	9,045
Eastern Anatolia R.	Ağrı	6,915	2,270	32.19%	2,136	6,002
	Ardahan	5,233	813	15.71%	2,133	4,423
	Bingöl	4,758	1,570	32.15%	2,261	5,656
	Bitlis	6,117	1,684	27.22%	2,309	5,414
	Elazığ	27,355	6,901	25.45%	1,925	4,791
	Erzincan	9,959	2,077	20.24%	2,250	5,294
	Erzurum	39,349	12,089	30.82%	2,012	4,432
	Hakkari	1,710	510	29.88%	3,231	8,951
	Iğdır	5,112	1,480	28.56%	2,654	8,397
	Kars	11,069	1,833	16.01%	1,889	4,655
	Malatya	36,342	8,386	23.82%	1,782	4,202
	Muş	5,812	1,647	28.38%	2,212	6,626
	Tunceli	1,458	283	19.24%	2,883	7,365
	Van	19,425	5,106	26.43%	2,431	6,507

Table 6.11: Descriptive statistics for cities - 2

Region	City	Policy Count	Claim Count	Claim Frequency	Claim Severity Average	Claim Severity Std. Dev.
Marmara R.	Balıkesir	95,914	21,903	22.84%	2,048	5,230
	Bilecik	13,766	2,533	18.07%	2,254	5,844
	Bursa	293,277	80,805	27.26%	1,848	4,190
	Çanakkale	44,596	9,232	20.47%	1,969	4,930
	Edirne	37,367	8,339	22.36%	1,820	4,539
	İstanbul	2,261,457	666,498	29.81%	2,121	5,094
	Kırklareli	28,811	5,979	20.21%	1,678	4,580
	Kocaeli	145,074	35,305	24.07%	2,307	4,917
	Sakarya	61,779	13,080	21.25%	2,372	5,314
	Tekirdağ	71,097	16,925	23.92%	1,870	4,701
	Yalova	19,645	4,194	21.07%	1,982	4,215
Mediterranean R.	Adana	192,177	49,214	25.63%	1,799	4,968
	Antalya	278,699	62,692	22.27%	2,111	4,971
	Burdur	15,468	2,447	15.71%	2,477	7,309
	Hatay	75,848	17,984	23.25%	1,864	4,664
	Isparta	24,153	5,067	20.45%	1,755	4,075
	Kahramanmaraş	32,989	9,213	27.24%	1,786	5,157
	Mersin	144,427	32,525	22.96%	2,079	5,416
	Osmaniye	19,766	5,057	25.63%	1,560	4,419
Southeastern Anatolia R.	Adıyaman	15,827	4,791	30.74%	1,701	5,255
	Batman	14,869	6,358	42.87%	1,610	4,102
	Diyarbakır	31,957	12,178	38.29%	1,996	5,372
	Gaziantep	94,887	28,267	29.35%	1,658	4,748
	Kilis	2,214	542	24.47%	1,105	2,626
	Mardin	16,696	5,493	32.68%	2,094	6,259
	Şanlıurfa	28,120	8,437	30.67%	2,049	6,248
	Siirt	6,639	2,626	39.36%	1,969	5,044
	Şırnak	5,782	2,452	42.15%	1,973	6,271
Total		7,063,703	1,860,103	26.48%	2,010	4,915

6.4 Groups According to Number of Claims and Regions

While the premium is calculated with the Bühlmann and Bühlmann-Straub credibility methods, the first two grouping approaches will be based on the number of claims and geographical regions of cities.

First, the groups are created according to the number of claims, while the cities are divided into groups from 3 to 10 after ordering from largest to smallest. If the cities cannot be divided equally into groups, ungrouped cities are added to the group with the lowest number of claims. The main purpose of always leaving the ungrouped cities in the last group is to increase the exposure of the group with the lowest exposure. Secondly, cities are divided into 7 groups geographically.

Table B.1 and Table B.2 show how cities are grouped according to the number of claims. For regional grouping, geographical regions are used and created 7 groups. These groups are the same as Table 6.10 and 6.11.

6.5 K-Means Clustering Application for Grouping and Results

In this section, it is explained how the k values are selected in the k-means clustering method applied to the cities in the MTPL dataset of Türkiye, which methods are applied while selecting, and which clusters the cities are divided into.

For the selection of the number of clusters, elbow and silhouette methods are applied to the dataset. K-means algorithm works with numerical values as input. Therefore, the claim severity average and the number of claims from 2009 to 2010 are used as input on an accident year-quarter basis.

First, the elbow method is used to determine the optimal number of k. In the graph of the elbow method, the points on the graph line with elbow-like decreases show the values that can be selected for k. In the MTPL dataset, the optimal number of clusters k can be 3, 6, and 9 as it is seen in Figure 6.2.

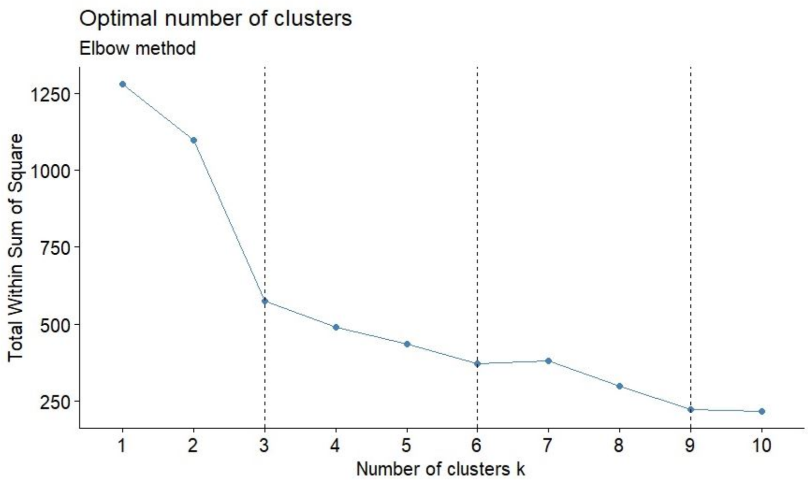

Figure 6.2: Elbow method for the number of clusters

Then, the silhouette method is applied to the dataset for the selection of the number of clusters. The value that maximizes the average silhouette width can be selected as the optimum number of clusters k for the k-means clustering algorithm. According to the silhouette method, Figure 6.3 shows the optimal number of clusters for the MTPL data is 3.

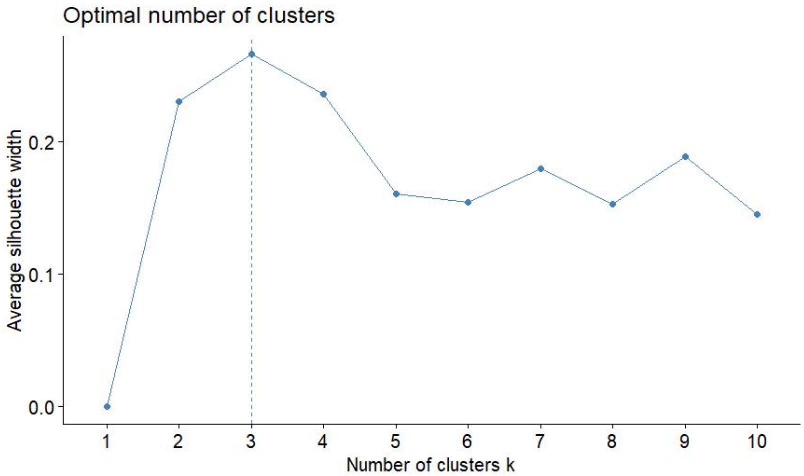

Figure 6.3: Silhouette method for the number of clusters

In the light of elbow and silhouette methods, the MTPL dataset is divided into 3, 6, and 9 clusters, respectively. After that, premiums are calculated by Bühlmann and Bühlmann-Straub method for cities in these groups separately. In the end, results are compared to select the best credibility method and the number of clusters.

The cluster plot is shown for $k = 3$ in Figure 6.4.

Figure 6.4: Number of clusters $k = 3$

As can be seen in Figure 6.4 with the green bullet İstanbul is clustered separately from other cities. The reason for that is the number of claims is sharply higher than in other cities.

When k is selected 6, the cluster plot is shown in Figure 6.5. As seen in Figure 6.5, as the number of clusters increases, the separation of those with high and low numbers of claims also increases. Tunceli and Hakkari which are shown in Figure 6.5 with 4 and 6 are with the lowest number of claims in the dataset. On the other hand, cities with high damage numbers such as İstanbul, Ankara, and Izmir also distinct from the other cities.

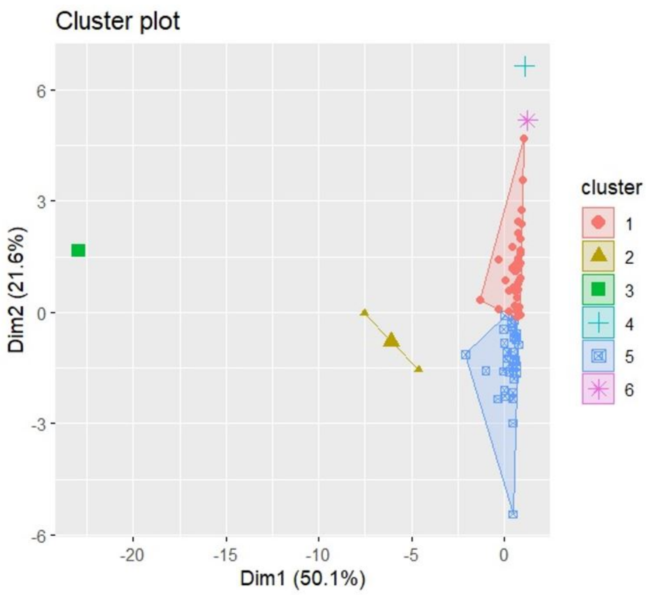

Figure 6.5: Number of clusters $k = 6$

For $k = 9$, the cluster plot is given in Figure 6.6. It is again seen that the distinction of the observations is directly proportional to the number of clusters. When $k = 9$ is selected, the majority of observations gather in main clusters. However, İstanbul, Ankara, Tunceli, İzmir, and Hakkari are still separated from the other observations by their characteristics.

Figure 6.6: Number of clusters $k = 9$

To conclude, premiums are calculated group by group in the light of clusters from the k-means algorithm. The lists of grouped cities are given in Table B.3 and Table B.4.

6.6 Groups Created by Hierarchical Clustering Method

In this section, clustering is applied by the hierarchical clustering method. The feasible number of clusters is examined by dendrograms. As a result of dendrograms, cities are split into 5 groups from 2 to 7. Lists of grouped cities are represented in Table B.3 and Table B.4. Moreover, the dendrograms are given in Appendix C.

When k is selected 2, only İstanbul is clustered separately. It is also seen in Figure C.1.

When k is selected 3, although there is a huge difference between the number of claims for Ankara, İzmir, and Kilis, interestingly, Kilis is clustered with these cities. It is shown in Figure C.2.

When other dendrograms are examined in Appendix C, it is seen that when k increases, cities with a low number of claims, such as Tunceli and Hakkari are grouped in different clusters by hierarchical clustering.

6.7 Comparison of Bühlmann and Bühlmann-Straub Results

Credibility factors and updated premiums are given for each city and the grouping method by tables in Appendix D. Results are compared with the 2011 Q1 (accident year-quarter) results which are given in Table D.13.

First, a comparison is made between the Bühlmann and Bühlmann-Straub estimates using the premiums calculated for the cities and the 2011 Q1 values. As a benchmark, the difference between the calculated premiums and the target premium (2011 Q1) is considered.

Table 6.12: Number of cities more approximate to target premium

Grouping Method	Group	Bühlmann	Bühlmann-Straub	Total
Number of Claims	$k=3$	14	67	81
	$k=4$	18	63	81
	$k=5$	21	60	81
	$k=6$	17	64	81
	$k=7$	20	61	81
	$k=8$	15	66	81
	$k=9$	19	62	81
	$k=10$	15	66	81
Region	$k=7$	17	64	81
K-means	$k=3$	14	67	81
	$k=6$	15	66	81
	$k=9$	13	68	81
Hierarchical	$k=2$	15	66	81
	$k=3$	15	66	81
	$k=4$	16	65	81
	$k=5$	16	65	81
	$k=6$	15	66	81
	$k=7$	18	63	81
Total		292	1,166	1,458

As can be seen in Table 6.12, results closer to the targeted premiums are obtained with the Bühlmann-Straub method in more than 60 cities in each grouping. Moreover, the total results show that Bühlmann-Straub predicts closer to the target in 1166 out of 1458 computations. We can conclude from Table 6.12 that in the MTPL dataset Bühlmann-Straub method estimates better than the Bühlmann method. However, it is not possible to comment on the performance of groups from this table. Another conclusion that can be seen from this table is that changing the number of groups does not create a significant improvement in the number of cities estimated closer to the target in the Bühlmann or Bühlmann-Straub model. A linear relationship cannot be seen between the number of groups and results.

Table 6.13: Credibility factor in accordance with credibility factor banding

		Bühlmann				Bühlmann-Straub			
Grouping Method	Group	0 - 0,25	0.25- 0.5	0.50 - 0.75	0.75 - 1	0 - 0,25	0.25- 0.5	0.50 - 0.75	0.75 - 1
Number of Claims	$k = 3$	0	54	27	0	3	26	43	9
	$k = 4$	0	61	0	20	0	19	56	6
	$k = 5$	0	33	32	16	0	30	31	20
	$k = 6$	13	29	26	13	13	25	14	29
	$k = 7$	11	26	33	11	0	21	48	12
	$k = 8$	20	21	20	20	9	24	23	25
	$k = 9$	18	27	18	18	9	14	39	19
	$k = 10$	16	17	32	16	15	13	27	26
Region	$k = 7$	14	8	59	0	26	15	27	13
K-means	$k = 3$	51	30	0	0	44	20	12	5
	$k = 6$	79	0	0	2	48	19	12	2
	$k = 9$	75	4	0	2	78	1	1	1
Hierarchical	$k = 2$	1	80	0	0	14	22	25	20
	$k = 3$	1	77	0	3	7	22	28	24
	$k = 4$	2	76	0	3	7	23	27	24
	$k = 5$	3	75	0	3	8	23	26	24
	$k = 6$	4	75	0	2	8	23	27	23
	$k = 7$	6	73	0	2	10	21	27	23

When the clustering results are examined in general, the 4 groups to be selected as the best group are those divided into 3, 4, 5, and 7 the groups separated according to the number of claims for Bühlmann-Straub credibility. This result is derived from Table 6.13 in light of only the credibility factors bands. The most striking feature of these groups at first glance is that there is no credibility factor less than 0.25. On the other hand, in hierarchical clustering, groups in which from $k = 3$ to $k = 6$ can be selected as the best according to the distribution of credibility factors. In these groups, the number of cities is higher in credibility factor bandings which are greater than 0.5 6.13. The selection of the most suitable clustering for the dataset will be improper without the prediction of premium being included in the analysis.

Observations grouped with k-means clustering show quite low performance with regard to count of cities that have $0 - 0.25$ credibility factor results. To mention that the credibility factor is 0, in Bühlmann, the city directly takes its average as the premium estimate. In the Bühlmann-Straub method, the weighted average of the group is accepted as the estimation for that city. In other words, no information is taken from past data. For this reason, comparing the results with their averages, especially for cities with a large number of groups and low exposure, will lead to wrong inferences.

The number of cities with a credibility factor of 0 is given in Table 6.14. There are 21 and 43 cities that have 0 credibility factor when $k = 9$ in Bühlmann and Bühlmann-Straub, respectively. When the number of groups increases, exposure shows a decrease for groups and it causes low reliability of the results for the less populated cities.

Table 6.14: Number of cities with 0 credibility factor

Grouping Method	Group	Bühlmann	Bühlmann-Straub
Number of Claims	$k = 3$	0	0
	$k = 4$	0	0
	$k = 5$	0	0
	$k = 6$	0	0
	$k = 7$	0	0
	$k = 8$	10	0
	$k = 9$	9	0
	$k = 10$	8	0
Region	$k = 7$	14	13
K-means	$k = 3$	1	1
	$k = 6$	3	3
	$k = 9$	21	43
Hierarchical	$k = 2$	1	1
	$k = 3$	1	1
	$k = 4$	2	2
	$k = 5$	3	3
	$k = 6$	4	4
	$k = 7$	6	6

When no insurance claim history based on cities is included in the grouping methodology, administrational region separation is used for this methodology. In this method, we have pre-determined seven regions for credibility modeling. One of the basic disadvantages of this methodology, cities can be grouped without using exposure information so cities with highly scattered exposures can coincide with the same group which causes the reliability of this methodology can be questionable. In terms of credibility factors, the geographical group has the number of 14 and 13 credibility factors with 0 for Bühlmann and Bühlmann-Straub, respectively in the MTPL dataset.

Considering the predictive power of groups based on premium estimation is investigated for each city. First, the difference between the target premium (2011 Q1) and the estimated premium will be taken as a measure of success for the groups in the premium estimates. These differences will be aggregated on a city basis and the results will be examined. Tables showing differences for all cities and grouping methods are shown in Appendix D.3, D.4,D.7and, D.8.

Table 6.15 shows once again that the Bühlmann model predicts better than the Bühlmann-Straub model. Between Bühlmann-Straub models, the best clustering method seems the group created with $k = 2$ by the hierarchical clustering because its difference is minimum. K-means algorithm with $k = 9$ which has maximum value can be selected as the worst clustering method from the table.

Table 6.15: Aggregation for premium differences

Grouping Method	Group	Bülhmann	Bühlmann -Straub
Number of Claims	$k = 3$	35,799	26,752
	$k = 4$	35,318	27,144
	$k = 5$	35,758	27,430
	$k = 6$	35,440	27,017
	$k = 7$	35,737	27,302
	$k = 8$	36,042	27,474
	$k = 9$	35,890	27,597
	$k = 10$	35,499	27,338
Region	$k = 7$	36,160	26,404
K-means	$k = 3$	36,116	26,824
	$k = 6$	35,397	26,365
	$k = 9$	37,302	28,484
Hierarchical	$k = 2$	36,946	25,895
	$k = 3$	36,716	25,986
	$k = 4$	36,918	26,870
	$k = 5$	36,940	27,286
	$k = 6$	36,940	26,585
	$k = 7$	37,423	27,478

The written policy numbers of İstanbul and Tunceli are 2,261,457 and 1,458, respectively. However, in Table 6.15 the impact of İstanbul and Tunceli is taken as the same in terms of premium difference. This approach may be used for balanced datasets, however, in the MTPL dataset, the percentage of written policy in İstanbul is 32.10%. Therefore, this calculation will be given as a weighted loss by the number of policy considerations.

Table 6.16: Weighted loss results

Grouping Method	Group	Bülhmann	Bühlmann -Straub
Number of Claims	$k = 3$	260.769	172.659
	$k = 4$	251.109	173.397
	$k = 5$	255.702	174.645
	$k = 6$	261.838	173.857
	$k = 7$	260.317	174.659
	$k = 8$	258.020	175.510
	$k = 9$	259.516	176.406
	$k = 10$	261.009	175.881
Region	$k = 7$	288.212	179.622
K-means	$k = 3$	285.028	172.703
	$k = 6$	288.073	173.737
	$k = 9$	266.032	173.459
Hierarchical	$k = 2$	320.338	172.148
	$k = 3$	297.112	171.573
	$k = 4$	293.324	171.619
	$k = 5$	290.479	171.544
	$k = 6$	293.560	171.483
	$k = 7$	292.868	172.170

In Table 6.16, results calculated with the weight of each city are shown in the dataset. In this table, it can be seen that the minimum difference belongs to the group with $k = 6$ in accordance with the hierarchical clustering for the Bühlmann-Straub method. Moreover, when differences are examined, it is seen that hierarchical clustering is the most suitable method for the dataset because their differences are lower than others in the Bühlmann-Straub method.

To sum up, when credibility factors banding result, the number of 0 credibility factors, and weighted loss prediction are considered together, it can be said that 6 groups by the hierarchical clustering can be the best clustering method for this dataset. The distribution of cities with this group is shown in Figure A.4 on the Türkiye map. In Table 6.17, descriptive statistics for this group are given by each cluster.

Table 6.17: Hierarchical clustering $k = 6$ - descriptive statistics for each group

Groups	1st Quarter	Median	Mean	3rd Quarter	Maximum	Std. Dev.	CoV
Group 1	401	826	2,024	1,826	649,442	5,039	2.4896
Group 2	317	654	1,924	1,583	222,194	4,823	2.5068
Group 3	299	590	1,879	1,477	232,607	4,975	2.6477
Group 4	296	583	2,010	1,510	422,970	5,553	2.7627
Group 5	289	581	2,210	1,603	226,155	6,351	2.8738
Group 6	285	557	2,198	1,593	259,395	6,228	2.8335

CHAPTER 7

CONCLUSION

Pricing is one of the most critical issues in the insurance industry. It is also of great importance to understand and predetermine the strategies to be followed, especially in insurance branches governed by regulations. In Türkiye, MTPL insurance is one of the instances of these kinds of insurance products. When it is examined the writing capacity in MTPL insurance, almost all non-life insurance companies in Türkiye write policies in this branch and MTPL insurance is one of the highest premium production items for Türkiye insurance sector.

Calculations in the dataset used are made on the basis of accident year-period. The premiums predicted by using the quarters of 2009 and 2010 are compared with the results of the first quarter of 2011.

Within the scope of this study, credibility methods are used for premium calculation in MTPL by using past claim averages and the number of claims and investigating the best group number with different scenarios. Four methodologies are chosen to group the dataset. The first methodology is that the cities are sorted from the largest to the smallest according to the total number of claims, and they are divided into 3 to 10 groups. Secondly, cities are divided into 7 groups according to just the geographical regions. Thirdly, the k-means clustering algorithm is applied, and the dataset is divided into 3, 6, and, 9 groups in the light of the elbow and silhouette method. Lastly, the hierarchical clustering method is used for clustering, and cities are divided into 2 to 7 groups according to the height indicators of the dendrogram. To sum up, in this thesis 18 grouping methods are applied to the dataset separately.

As a first comparison, the Bühlmann-Straub model estimated results closer to the 2011 first quarter in each grouping than the Bühlmann model. Afterward, it is investigated by which method the best clustering is obtained in the Bühlmann-Straub credibility model. It is seen that in the comparisons made by considering different metrics together, the method in which the best clustering for the Türkiye MTPL dataset is the hierarchical clustering with 6 groups.

This study can be used for reasons that premium calculation for MTPL insurance and cap premium prediction by insurance companies and Insurance and Private Pension Regulation and Supervision Agency (PPRSA).

Last but not least, in further studies Bühlmann and Bühlmann-Straub credibility methods can be applied with the policyholder or car segmentation to have more precise groups for premium predictions.

REFERENCES

[1] 2021 tsb sector report, https://tsb.org.tr/media/attachments/TSB_SEKTOR_TR21_2807.pdf, Accessed: 2022-12-31.

[2] Coverage limits, https://www.mevzuat.gov.tr/anasayfa/MevzuatFihristDetayIframe?MevzuatTur=7&MevzuatNo=11444&MevzuatTertip=5, Accessed: 2022-12-31.

[3] Cpi indicators, https://data.tuik.gov.tr/Bulten/Index?p=Tuketici-Fiyat-Endeksi-Ekim-2022-45799#:~:text=T%C3%9CFE'deki%20(2003%3D100,%65%2C26%20olarak%20ger%C3%A7ekle%C5%9Fti.&text=Bir%20%C3%B6nceki%20y%C4%B1l%C4%B1n%20ayn%C4%B1%20ay%C4%B1na%20g%C3%B6re%20en%20az%20art%C4%B1%C5%9F%20g%C3%B6steren,33%2C48%20ile%20haberle%C5%9Fme%20oldu., Accessed: 2022-12-31.

[4] Highway law, no. 2918, https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=2918&MevzuatTur=1&MevzuatTertip=5, Accessed: 2022-12-31.

[5] Mtpl cap premium circular, https://www.resmigazete.gov.tr/eskiler/2017/05/20170503-8.htm, Accessed: 2022-12-31.

[6] Mtpl premium increases, https://www.resmigazete.gov.tr/eskiler/2022/08/20220812.pdf, Accessed: 2022-12-31.

[7] Türkiye insurance sector results, https://www.tsb.org.tr/tr/istatistikler, Accessed: 2022-12-31.

[8] H. Bühlmann, Experience rating and credibility, ASTIN Bulletin: The Journal of the IAA, 4(3), pp. 199–207, 1967.

[9] H. Bühlmann and A. Gisler, *A course in credibility theory and its applications*, volume 317, Springer, 2005.

[10] H. Buhlmann and E. Straub, Credibility for loss ratios, Actuarial Research Clearing House, 2, 1972.

[11] S. E. Bülbül and K. B. Baykal, Optimal bonus-malus system design in motor third-party liability insurance in turkey: Negative binomial model, International Journal of Economics and Finance, 8(8), pp. 205–211, 2016.

[12] Q. C. Chukwudum, Credibility premium estimation of insurance claims in nigeria, 2018.

[13] L. Diao and C. Weng, Regression tree credibility model, North American Actuarial Journal, 23(2), pp. 169–196, 2019.

[14] H. Dornheim and V. Brazauskas, Robust–efficient credibility models with heavy-tailed claims: A mixed linear models perspective, Insurance: Mathematics and Economics, 48(1), pp. 72–84, 2011.

[15] C. Dutang, V. Goulet, and M. Pigeon, actuar: An r package for actuarial science, Journal of Statistical software, 25, pp. 1–37, 2008.

[16] J. Gareth, W. Daniela, H. Trevor, and T. Robert, *An introduction to statistical learning: with applications in R*, Spinger, 2013.

[17] E. Gómez-Déniz, A generalization of the credibility theory obtained by using the weighted balanced loss function, Insurance: Mathematics and Economics, 42(2), pp. 850–854, 2008.

[18] C. Hachemeister, Credibility for regression models with application to trend (reprint), Credibility: Theory and Applications. Edited by P. Kahn. New York: Academic Press, Inc, pp. 307–48, 1975.

[19] A. Hassan Zadeh and D. A. Stanford, Bayesian and bühlmann credibility for phase-type distributions with a univariate risk parameter, Scandinavian Actuarial Journal, 2016(4), pp. 338–355, 2016.

[20] T. N. Herzog, *Introduction to Credibility Theory*, ACTEX Academic Series, ACTEX Publications, Inc., 2010, ISBN 978-1-56698-764-6.

[21] R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit, *Modern actuarial risk theory: using R*, volume 128, Springer Science & Business Media, 2008.

[22] J. H. Kim and Y. Jeon, Credibility theory based on trimming, Insurance: Mathematics and Economics, 53(1), pp. 36–47, 2013.

[23] B. Linda and J. Kubanová, Credibility premium calculation in motor third-party liability insurance, in *WSEAS conference Advances in Mathematical and Computational Methods, Malta*, pp. 259–263, 2012.

[24] A. Lotsi, F. O. Mettle, and P. K. Adjorlolo, Application of bühlmanns-straub credibility theory in determining the effect of frequency-severity on credibility premium estimation, ADRRI Journal of Physical and Natural Sciences, 3(2), pp. 1–24, 2019.

[25] A. H. Mowbray, How extensive a payroll exposure is necessary to give a dependable pure premium, in *Proceedings of the Casualty Actuarial Society*, volume 1, pp. 24–30, 1914.

[26] A. T. P. Najafabadi, A new approach to the credibility formula, Insurance: Mathematics and Economics, 46(2), pp. 334–338, 2010.

[27] J. A. Nelder and R. J. Verrall, Credibility theory and generalized linear models, ASTIN Bulletin: The Journal of the IAA, 27(1), pp. 71–82, 1997.

[28] S. Pitrebois, M. Denuit, and J.-F. Walhin, An actuarial analysis of the french bonus-malus system, Scandinavian Actuarial Journal, 2006(5), pp. 247–264, 2006.

[29] G. Pitselis, Pure robust versus robust portfolio unbiased—credibility and asymptotic optimality, Insurance: Mathematics and Economics, 52(2), pp. 391–403, 2013.

[30] G. Pitselis, Quantile credibility models, Insurance: Mathematics and Economics, 52(3), pp. 477–489, 2013.

[31] G. Pitselis, Credible risk measures with applications in actuarial sciences and finance, Insurance: Mathematics and Economics, 70, pp. 373–386, 2016.

[32] G. Pitselis, Risk measures in a quantile regression credibility framework with fama/french data applications, Insurance: Mathematics and Economics, 74, pp. 122–134, 2017.

[33] P. Shi and L. Yang, Pair copula constructions for insurance experience rating, Journal of the American Statistical Association, 113(521), pp. 122–133, 2018.

[34] Y.-K. Tse, *Nonlife actuarial models: theory, methods and evaluation*, Cambridge University Press, 2009.

[35] H. R. Waters, *Credibility Theory*, Department of Actuarial Maths and Statistics, Heriot-Watt University, 1993.

[36] L. Wen, X. Wu, and X. Zhou, The credibility premiums for models with dependence induced by common effects, Insurance: Mathematics and Economics, 44(1), pp. 19–25, 2009.

[37] A. W. Whitney, Theory of experience rating, in *Proceedings of the Casualty Actuarial Society*, volume 4, pp. 274—-292, 1918.

[38] V. Winarta, M. Novita, and S. Nurrohmah, Multivariate bühlmann-straub credibility model for claim reserving, in *Journal of Physics: Conference Series*, volume 1725, p. 012026, IOP Publishing, 2021.

[39] J. Wu, *Advances in K-means clustering: a data mining thinking*, Springer Science & Business Media, 2012.

[40] Y. Yan and K.-S. Song, A general optimal approach to bühlmann credibility theory, Insurance: Mathematics and Economics, 104, pp. 262–282, 2022.

[41] B. Yıldırım Külekci and A. S. Selcuk-Kestel, Assessment of longevity risk: credibility approach, Journal of Applied Statistics, 48(13-15), pp. 2695–2713, 2021.

APPENDIX A

APPLIED INFLATION FACTORS AND MAPS OF TÜRKİYE

Table A.1: Applied inflation factors by claim year-month

Year Month	Inflation Changes	Cumulative Inflation	Inflation Loading Factor
2009-1	0.003	1.000	1.148
2009-2	-0.003	0.997	1.152
2009-3	0.011	1.008	1.139
2009-4	0.000	1.008	1.139
2009-5	0.006	1.014	1.132
2009-6	0.001	1.015	1.131
2009-7	0.003	1.018	1.128
2009-8	-0.003	1.015	1.131
2009-9	0.004	1.019	1.127
2009-10	0.024	1.043	1.100
2009-11	0.013	1.057	1.086
2009-12	0.005	1.062	1.081
2010-1	0.019	1.082	1.061
2010-2	0.015	1.098	1.046
2010-3	0.006	1.104	1.040
2010-4	0.006	1.110	1.034
2010-5	-0.004	1.107	1.037
2010-6	-0.006	1.100	1.043
2010-7	-0.005	1.095	1.048
2010-8	0.004	1.099	1.044
2010-9	0.012	1.113	1.031
2010-10	0.018	1.133	1.013
2010-11	0.000	1.134	1.013
2010-12	-0.003	1.130	1.016
2011-1	0.004	1.135	1.012
2011-2	0.007	1.143	1.004
2011-3	0.004	1.148	1.000

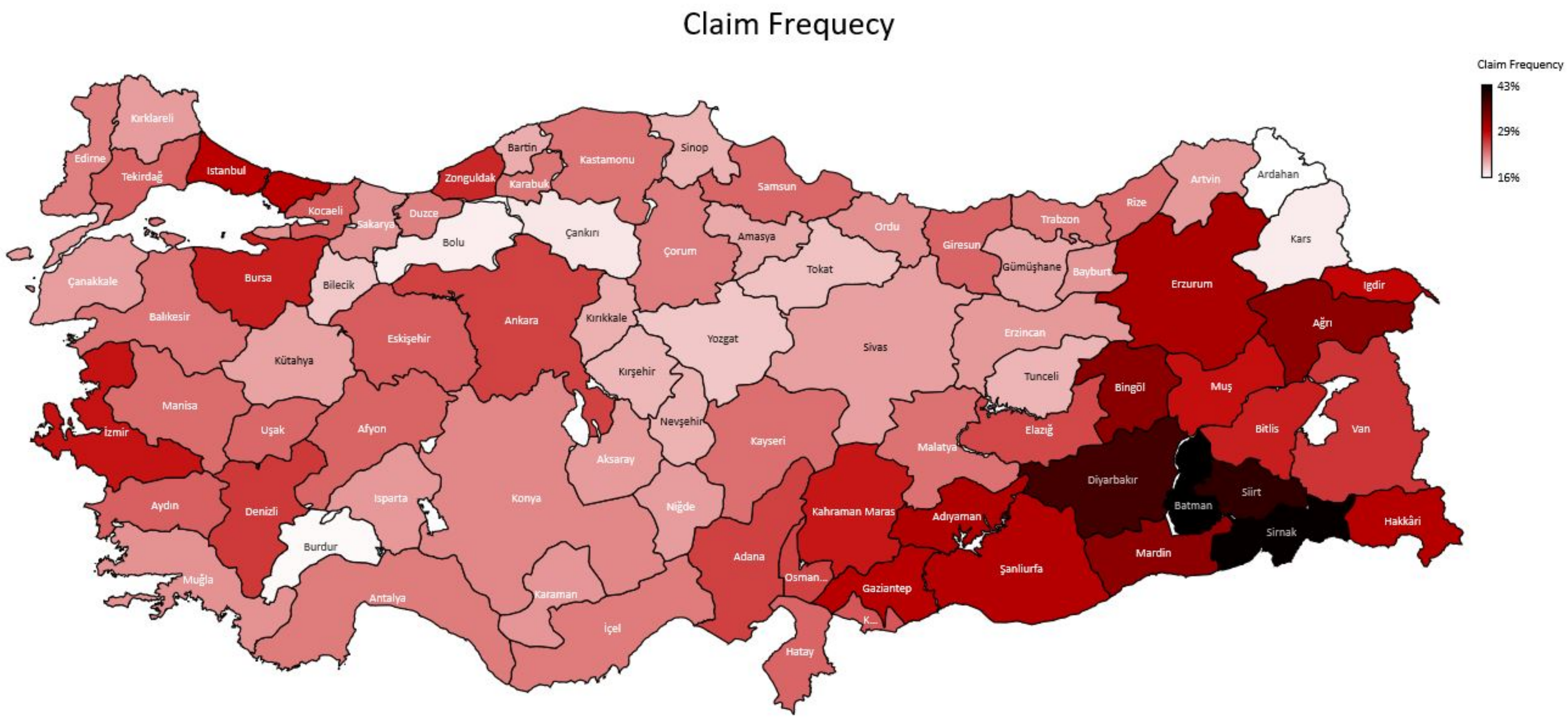

Figure A.1: Türkiye claim frequency distribution

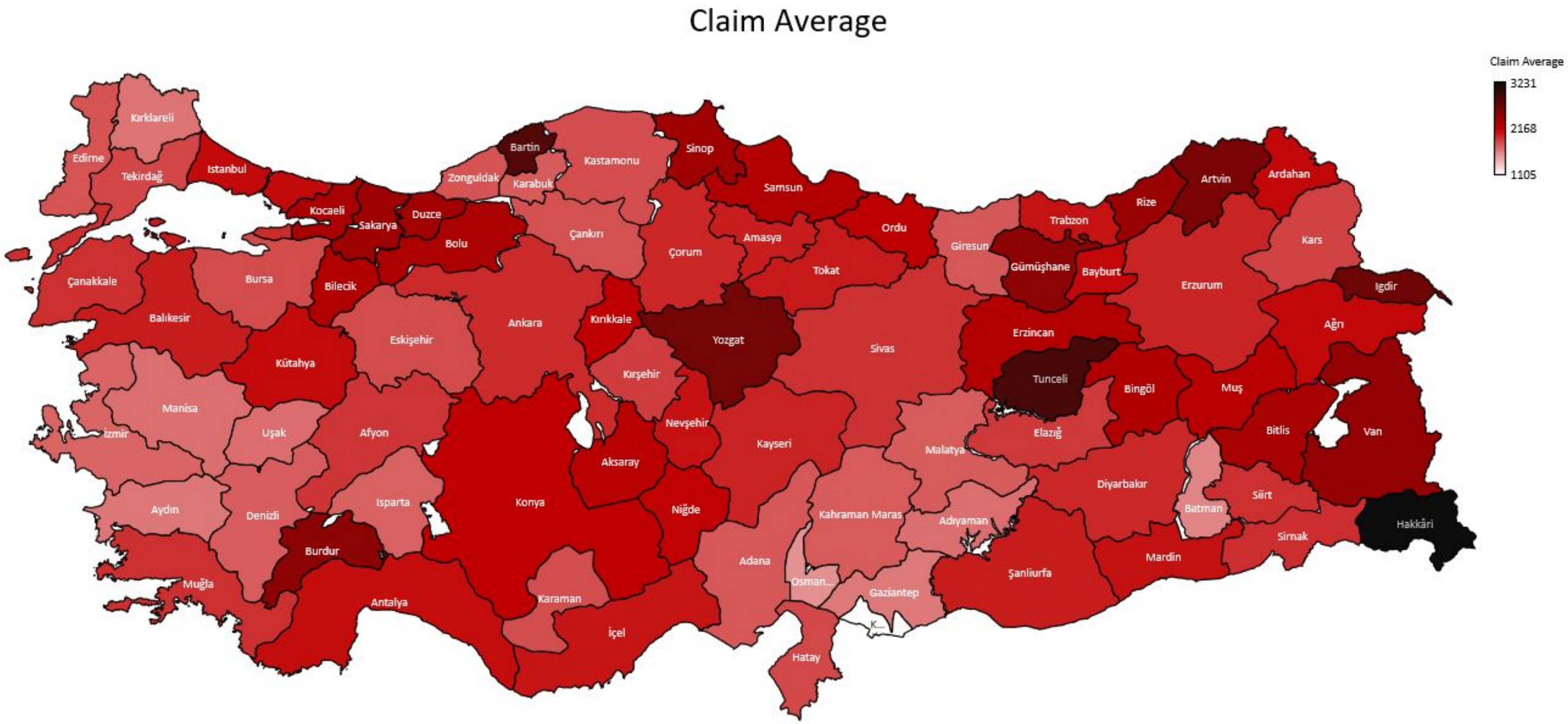

Figure A.2: Türkiye claim average distribution

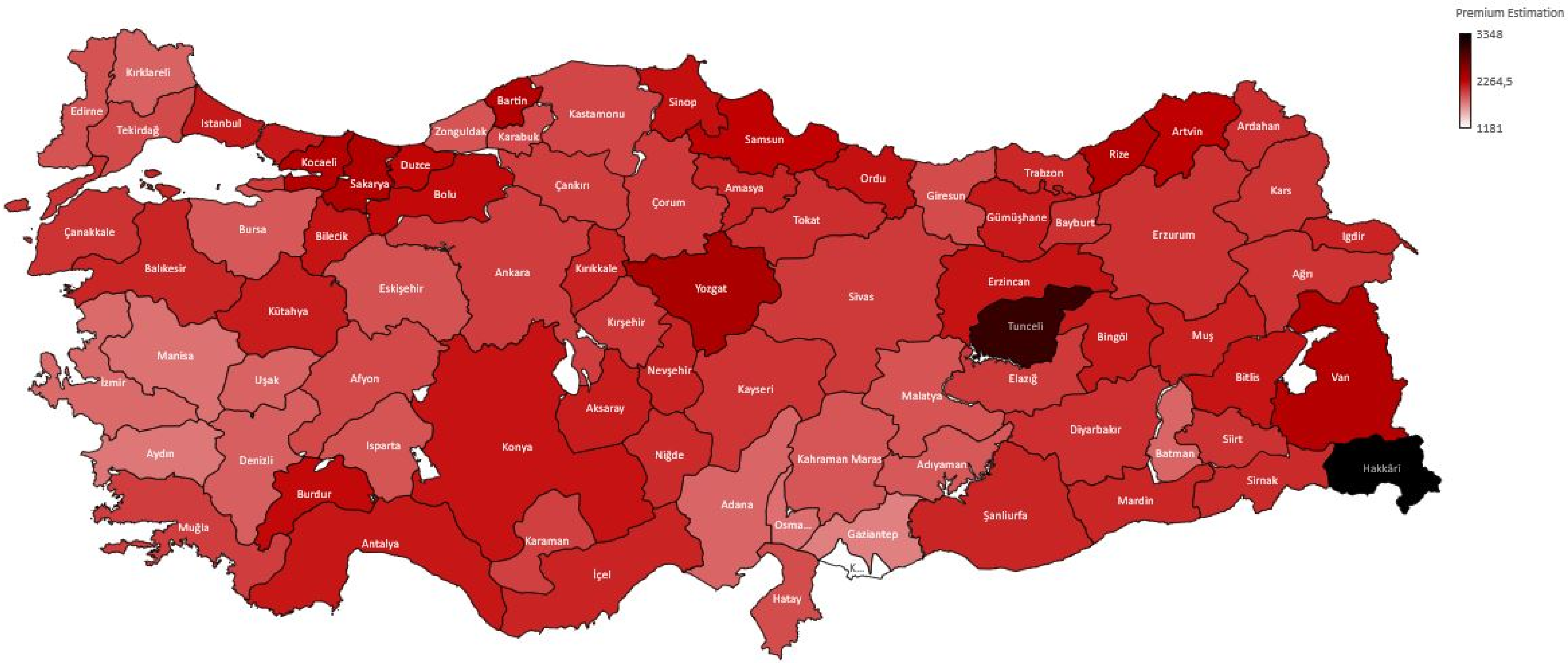

Figure A.3: Bühlmann-Straub premium map with hierarchical clustering 6 groups

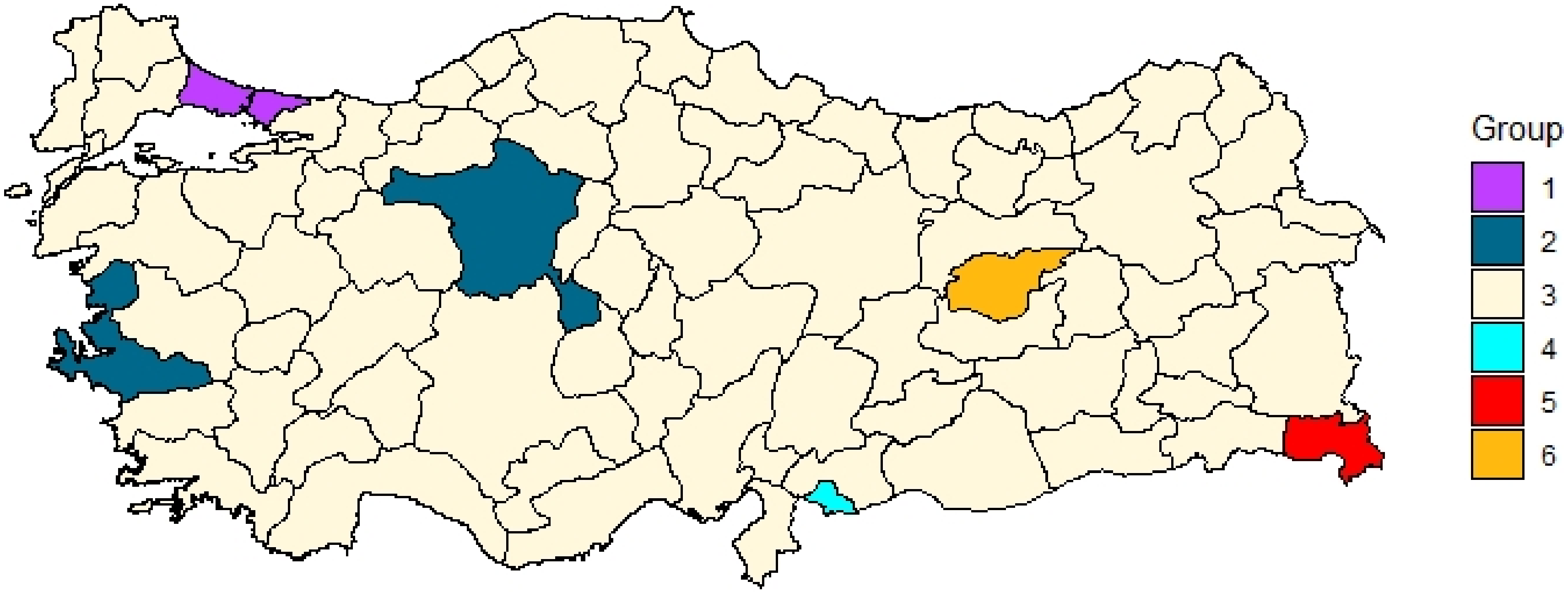

Figure A.4: The distribution of cities by the hierarchical clustering when $k = 6$

APPENDIX B

LIST OF CITY GROUPS BY CLUSTERING METHODS

Table B.1: City groups by the number of claims - 1

City	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$
İstanbul	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Ankara	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
İzmir	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Bursa	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Antalya	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Adana	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Kocaeli	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Mersin	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Gaziantep	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 2
Kayseri	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1	Group 2	Group 2
Samsun	Group 1	Group 1	Group 1	Group 1	Group 1	Group 2	Group 2	Group 2
Balıkesir	Group 1	Group 1	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2
Muğla	Group 1	Group 1	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2
Denizli	Group 1	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2	Group 2
Manisa	Group 1	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2	Group 2
Aydın	Group 1	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2	Group 2
Hatay	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2	Group 2	Group 3
Eskişehir	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2	Group 2	Group 3
Konya	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2	Group 3	Group 3
Tekirdağ	Group 1	Group 1	Group 2	Group 2	Group 2	Group 2	Group 3	Group 3
Trabzon	Group 1	Group 2	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3
Sakarya	Group 1	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3	Group 3
Zonguldak	Group 1	Group 2	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3
Erzurum	Group 1	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3	Group 3
Diyarbakır	Group 1	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3	Group 4
Ordu	Group 1	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3	Group 4
Çanakkale	Group 1	Group 2	Group 2	Group 3	Group 3	Group 3	Group 3	Group 4
Kahramanmaraş	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3	Group 4	Group 4
Malatya	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3	Group 4	Group 4
Şanlıurfa	Group 2	Group 2	Group 2	Group 3	Group 3	Group 3	Group 4	Group 4
Edirne	Group 2	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 4
Afyonkarahisar	Group 2	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 4
Çorum	Group 2	Group 2	Group 3	Group 3	Group 3	Group 4	Group 4	Group 5
Elazığ	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 4	Group 5
Giresun	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 4	Group 5
Sivas	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 4	Group 5
Batman	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 5	Group 5
Kırklareli	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 5	Group 5
Tokat	Group 2	Group 2	Group 3	Group 3	Group 4	Group 4	Group 5	Group 5
Uşak	Group 2	Group 2	Group 3	Group 4	Group 4	Group 4	Group 5	Group 5

Table B.2: City groups by the number of claims - 2

City	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$
Mardin	Group 2	Group 3	Group 3	Group 4	Group 4	Group 5	Group 5	Group 6
Van	Group 2	Group 3	Group 3	Group 4	Group 4	Group 5	Group 5	Group 6
Kütahya	Group 2	Group 3	Group 3	Group 4	Group 4	Group 5	Group 5	Group 6
Isparta	Group 2	Group 3	Group 3	Group 4	Group 4	Group 5	Group 5	Group 6
Düzce	Group 2	Group 3	Group 3	Group 4	Group 5	Group 5	Group 5	Group 6
Osmaniye	Group 2	Group 3	Group 3	Group 4	Group 5	Group 5	Group 6	Group 6
Adıyaman	Group 2	Group 3	Group 3	Group 4	Group 5	Group 5	Group 6	Group 6
Rize	Group 2	Group 3	Group 3	Group 4	Group 5	Group 5	Group 6	Group 6
Amasya	Group 2	Group 3	Group 4	Group 4	Group 5	Group 5	Group 6	Group 7
Kastamonu	Group 2	Group 3	Group 4	Group 4	Group 5	Group 5	Group 6	Group 7
Yalova	Group 2	Group 3	Group 4	Group 4	Group 5	Group 6	Group 6	Group 7
Nevşehir	Group 2	Group 3	Group 4	Group 5	Group 5	Group 6	Group 6	Group 7
Aksaray	Group 2	Group 3	Group 4	Group 4	Group 5	Group 6	Group 6	Group 7
Niğde	Group 2	Group 3	Group 4	Group 5	Group 5	Group 6	Group 6	Group 7
Bolu	Group 3	Group 3	Group 4	Group 5	Group 5	Group 6	Group 7	Group 7
Karabük	Group 3	Group 3	Group 4	Group 5	Group 6	Group 6	Group 7	Group 7
Siirt	Group 3	Group 3	Group 4	Group 5	Group 6	Group 6	Group 7	Group 8
Bilecik	Group 3	Group 3	Group 4	Group 5	Group 6	Group 6	Group 7	Group 8
Yozgat	Group 3	Group 3	Group 4	Group 5	Group 6	Group 6	Group 7	Group 8
Şırnak	Group 3	Group 3	Group 4	Group 5	Group 6	Group 6	Group 7	Group 8
Burdur	Group 3	Group 4	Group 4	Group 5	Group 6	Group 7	Group 7	Group 8
Ağrı	Group 3	Group 4	Group 4	Group 5	Group 6	Group 7	Group 7	Group 8
Artvin	Group 3	Group 4	Group 4	Group 5	Group 6	Group 7	Group 7	Group 8
Sinop	Group 3	Group 4	Group 4	Group 5	Group 6	Group 7	Group 8	Group 8
Erzincan	Group 3	Group 4	Group 5	Group 5	Group 6	Group 7	Group 8	Group 9
Kırşehir	Group 3	Group 4	Group 5	Group 6	Group 6	Group 7	Group 8	Group 9
Kars	Group 3	Group 4	Group 5	Group 6	Group 7	Group 7	Group 8	Group 9
Bitlis	Group 3	Group 4	Group 5	Group 6	Group 7	Group 7	Group 8	Group 9
Muş	Group 3	Group 4	Group 5	Group 6	Group 7	Group 7	Group 8	Group 9
Karaman	Group 3	Group 4	Group 5	Group 6	Group 7	Group 7	Group 8	Group 9
Bartın	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 8	Group 9
Bingöl	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 8	Group 9
Iğdır	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Çankırı	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Kırıkkale	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Gümüşhane	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Ardahan	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Kilis	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Bayburt	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Hakkari	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10
Tunceli	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10

Table B.3: City groups by region, k-means and hierarchical clustering - 1

	Region	K-means			Hierarchical					
City	$k = 7$	$k = 3$	$k = 6$	$k = 9$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
İstanbul	Group 1	Group 2	Group 3	Group 6	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Ankara	Group 2	Group 3	Group 2	Group 2	Group 2	Group 2	Group 2	Group 2	Group 2	Group 2
İzmir	Group 3	Group 3	Group 2	Group 2	Group 2	Group 2	Group 2	Group 2	Group 2	Group 2
Bursa	Group 1	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Antalya	Group 4	Group 3	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Adana	Group 4	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kocaeli	Group 1	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Mersin	Group 4	Group 3	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Gaziantep	Group 5	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kayseri	Group 2	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Samsun	Group 6	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Balıkesir	Group 1	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Muğla	Group 3	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Denizli	Group 3	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Manisa	Group 3	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Aydın	Group 3	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Hatay	Group 4	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Eskişehir	Group 2	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Konya	Group 2	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Tekirdağ	Group 1	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Trabzon	Group 6	Group 3	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Sakarya	Group 1	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Zonguldak	Group 6	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Erzurum	Group 7	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Diyarbakır	Group 5	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Ordu	Group 6	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Çanakkale	Group 1	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kahramanmaraş	Group 4	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Malatya	Group 7	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Şanlıurfa	Group 5	Group 3	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Edirne	Group 1	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Afyonkarahisar	Group 3	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Çorum	Group 6	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Elazığ	Group 7	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Giresun	Group 6	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Sivas	Group 2	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Batman	Group 5	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kırklareli	Group 1	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Tokat	Group 6	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Uşak	Group 3	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3

Table B.4: City groups by region, k-means and hierarchical clustering - 2

	Region	K-means			Hierarchical					
City	$k=7$	$k=3$	$k=6$	$k=9$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
Mardin	Group 5	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Van	Group 7	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kütahya	Group 3	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Isparta	Group 4	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Düzce	Group 6	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Osmaniye	Group 4	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Adıyaman	Group 5	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Rize	Group 6	Group 1	Group 1	Group 7	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Amasya	Group 6	Group 3	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kastamonu	Group 6	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Yalova	Group 1	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Nevşehir	Group 2	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Aksaray	Group 2	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Niğde	Group 2	Group 3	Group 1	Group 3	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Bolu	Group 6	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Karabük	Group 6	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Siirt	Group 5	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Bilecik	Group 1	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Yozgat	Group 2	Group 1	Group 1	Group 3	Group 2	Group 3	Group 3	Group 3	Group 3	Group 4
Şırnak	Group 5	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Burdur	Group 4	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Ağrı	Group 7	Group 3	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Artvin	Group 6	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Sinop	Group 6	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Erzincan	Group 7	Group 1	Group 1	Group 7	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kırşehir	Group 2	Group 3	Group 5	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kars	Group 7	Group 3	Group 5	Group 3	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Bitlis	Group 7	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Muş	Group 7	Group 1	Group 1	Group 7	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Karaman	Group 2	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Bartın	Group 6	Group 1	Group 1	Group 7	Group 2	Group 3	Group 3	Group 3	Group 3	Group 4
Bingöl	Group 7	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Iğdır	Group 7	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Çankırı	Group 2	Group 3	Group 5	Group 1	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kırıkkale	Group 2	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Gümüşhane	Group 6	Group 1	Group 1	Group 9	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Ardahan	Group 7	Group 3	Group 1	Group 3	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Kilis	Group 5	Group 3	Group 5	Group 1	Group 2	Group 2	Group 2	Group 2	Group 4	Group 5
Bayburt	Group 6	Group 1	Group 1	Group 4	Group 2	Group 3	Group 3	Group 3	Group 3	Group 3
Hakkari	Group 7	Group 1	Group 4	Group 8	Group 2	Group 3	Group 3	Group 4	Group 5	Group 6
Tunceli	Group 7	Group 1	Group 6	Group 5	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7

APPENDIX C

DENDROGRAM FIGURES FOR HIERARCHICAL CLUSTERING

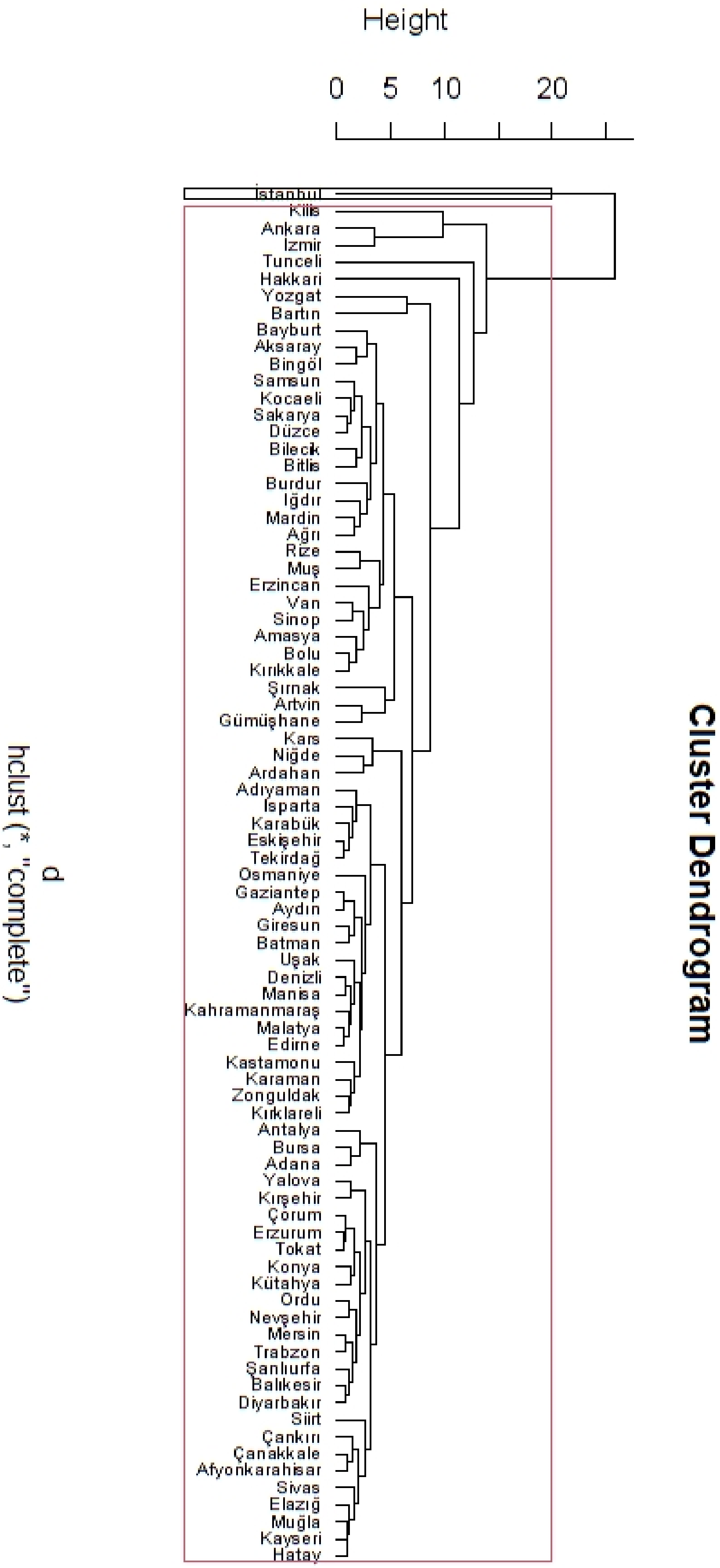

Figure C.1: Dendrogram for $k = 2$

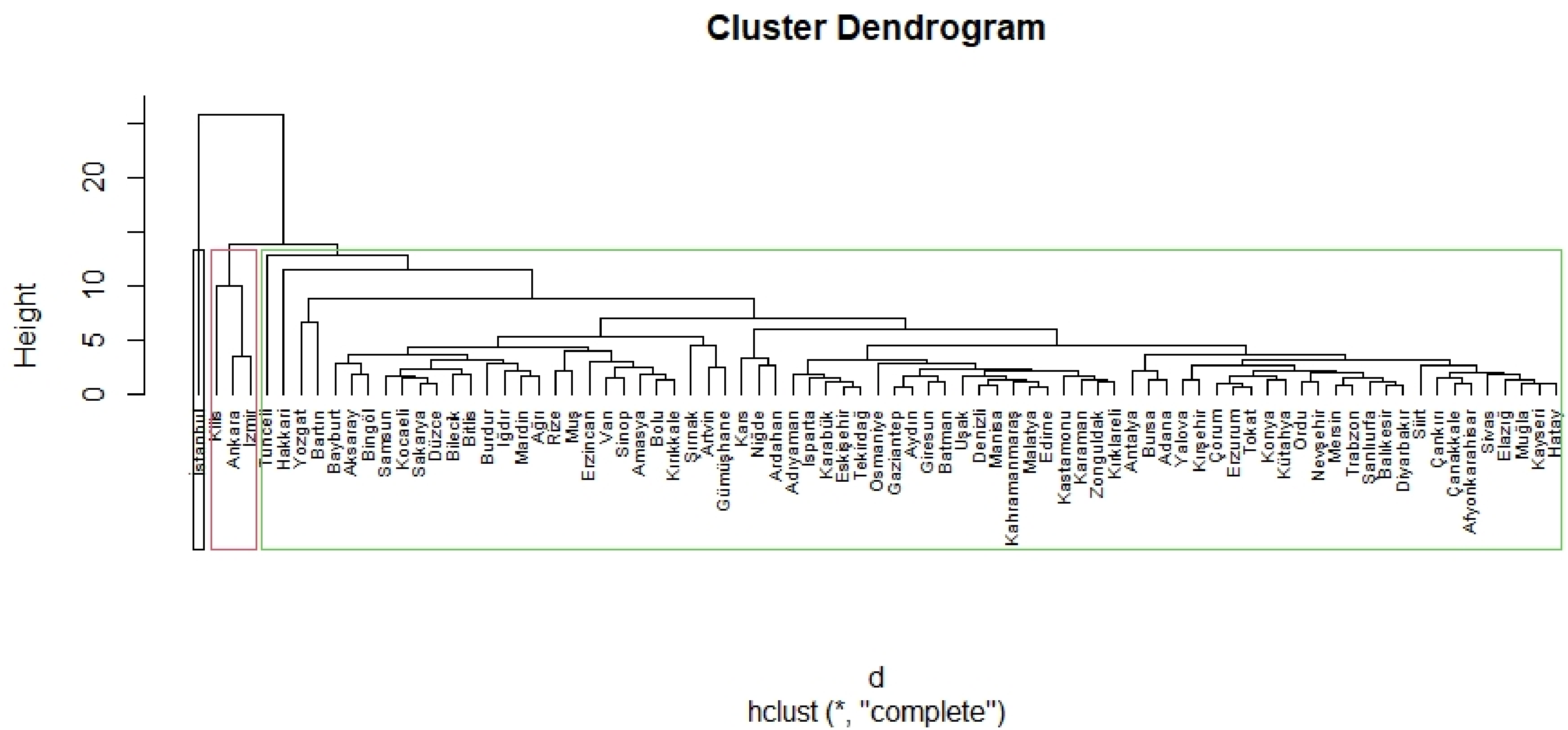

Figure C.2: Dendrogram for $k = 3$

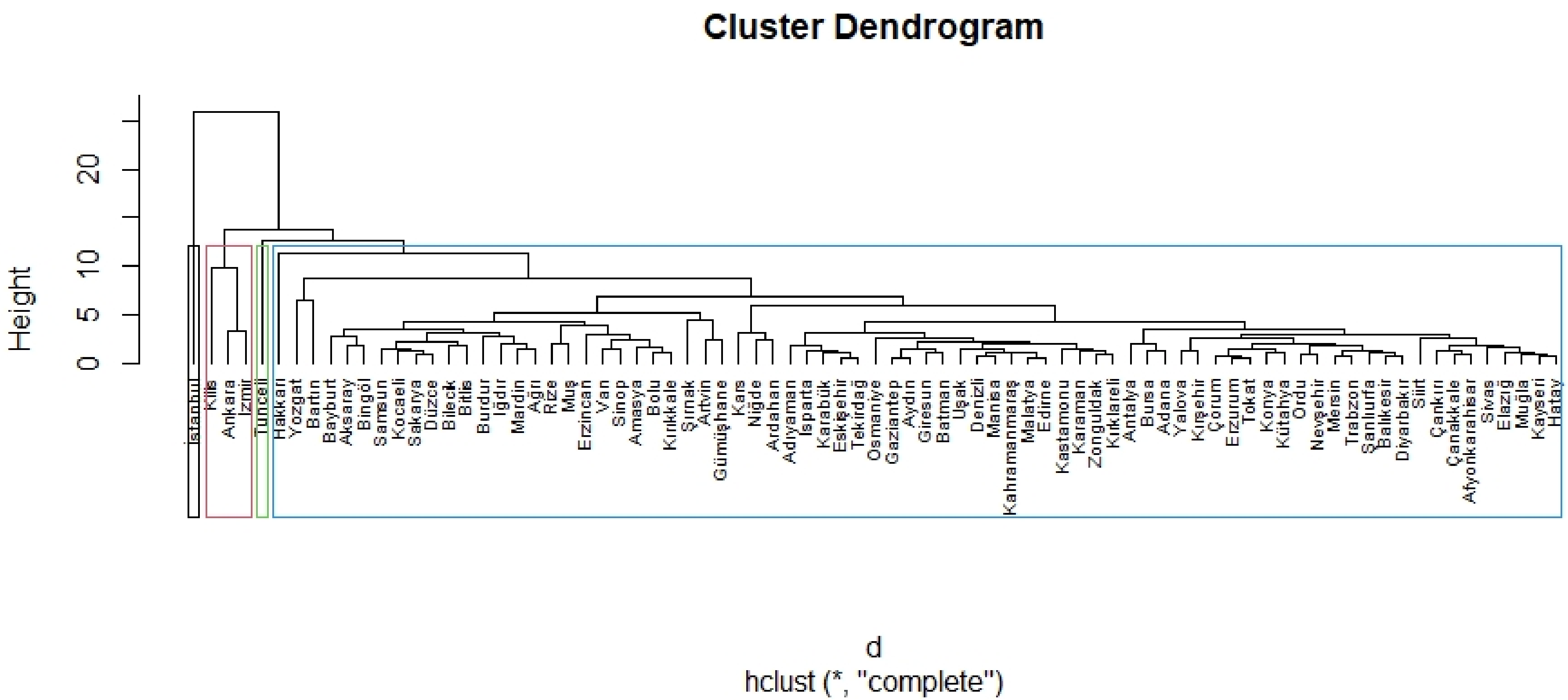

Figure C.3: Dendrogram for $k = 4$

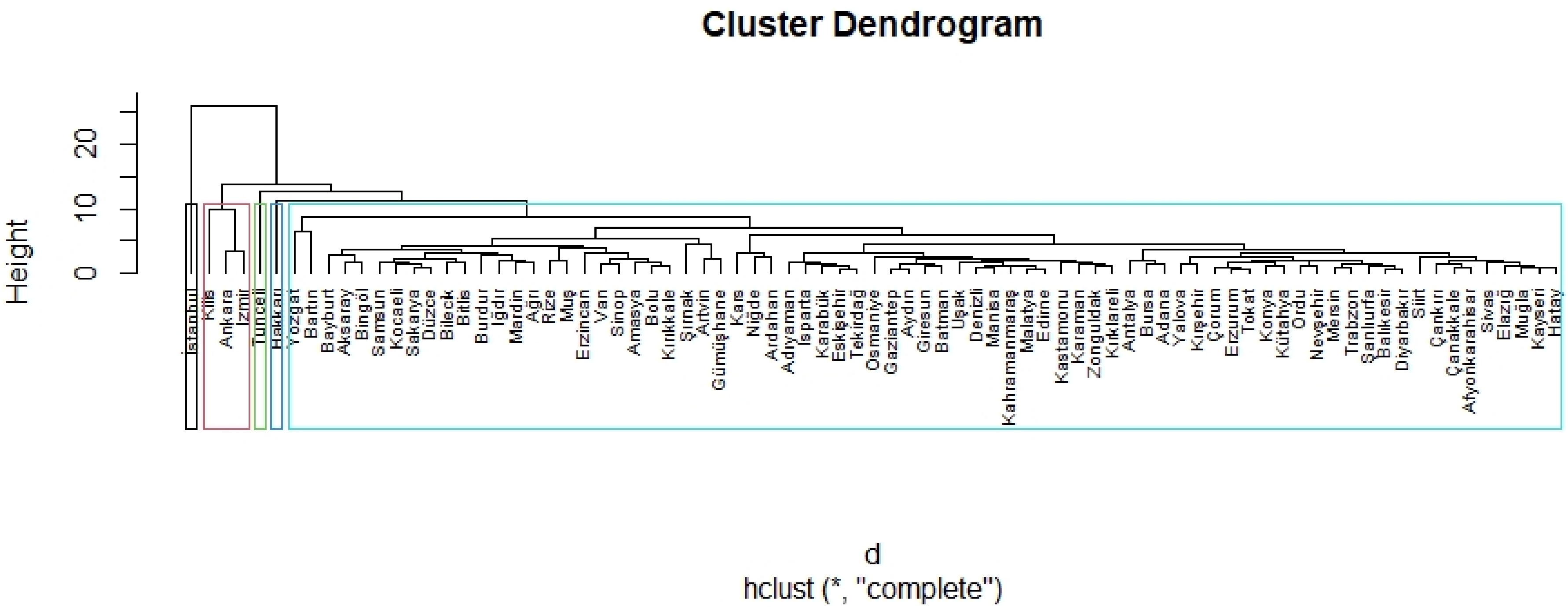

Figure C.4: Dendrogram for $k = 5$

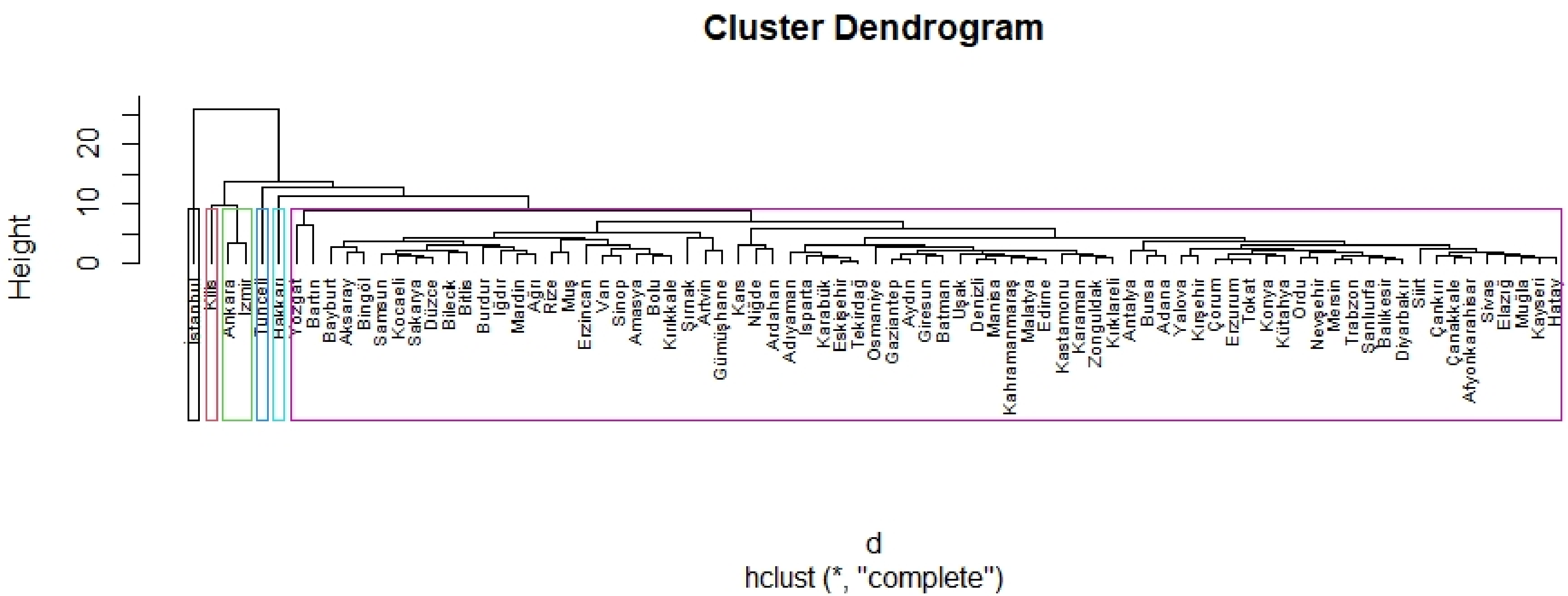

Figure C.5: Dendrogram for $k = 6$

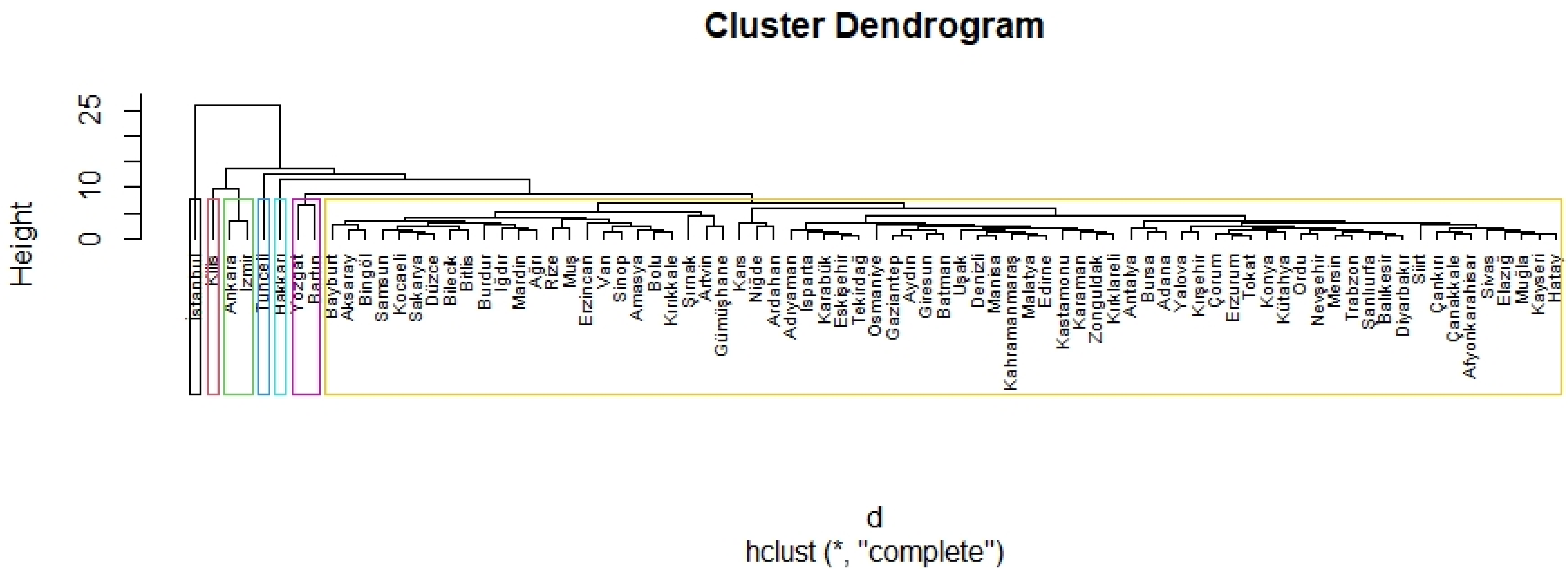

Figure C.6: Dendrogram for $k = 7$

APPENDIX D

BÜHLMANN AND BÜHLMANN-STRAUB CREDIBILITY RESULTS

Table D.1: Bühlmann premiums and credibility Factors - 1

City	$k=3$		$k=4$		$k=5$		$k=6$		$k=7$		$k=8$		$k=9$		$k=10$	
İstanbul	2,224	71.97%	2,221	77.55%	2,227	80.43%	2,233	76.02%	2,238	81.59%	2,233	81.90%	2,236	83.31%	2,240	81.53%
Ankara	2,119	71.97%	2,109	77.55%	2,111	80.43%	2,123	76.02%	2,120	81.59%	2,114	81.90%	2,115	83.31%	2,122	81.53%
İzmir	1,928	71.97%	1,903	77.55%	1,897	80.43%	1,921	76.02%	1,904	81.59%	1,897	81.90%	1,894	83.31%	1,906	81.53%
Bursa	1,994	71.97%	1,974	77.55%	1,971	80.43%	1,991	76.02%	1,978	81.59%	1,972	81.90%	1,970	83.31%	1,980	81.53%
Antalya	2,248	71.97%	2,247	77.55%	2,254	80.43%	2,259	76.02%	2,266	81.59%	2,260	81.90%	2,264	83.31%	2,267	81.53%
Adana	1,920	71.97%	1,895	77.55%	1,888	80.43%	1,913	76.02%	1,894	81.59%	1,888	81.90%	1,885	83.31%	1,896	81.53%
Kocaeli	2,367	71.97%	2,376	77.55%	2,388	80.43%	2,384	76.02%	2,401	81.59%	2,396	81.90%	2,402	83.31%	2,402	81.53%
Mersin	2,187	71.97%	2,182	77.55%	2,187	80.43%	2,195	76.02%	2,197	81.59%	2,191	81.90%	2,194	83.31%	2,199	81.53%
Gaziantep	1,862	71.97%	1,832	77.55%	1,823	80.43%	1,851	76.02%	1,829	81.59%	1,822	81.90%	1,817	83.31%	1,814	79.85%
Kayseri	2,080	71.97%	2,066	77.55%	2,067	80.43%	2,081	76.02%	2,075	81.59%	2,069	81.90%	2,057	76.35%	2,055	79.85%
Samsun	2,352	71.97%	2,359	77.55%	2,371	80.43%	2,368	76.02%	2,383	81.59%	2,346	75.14%	2,345	76.35%	2,357	79.85%
Balıkesir	2,225	71.97%	2,223	77.55%	2,230	80.43%	2,235	76.02%	2,173	60.07%	2,215	75.14%	2,212	76.35%	2,217	79.85%
Muğla	2,127	71.97%	2,117	77.55%	2,119	80.43%	2,131	76.02%	2,090	60.07%	2,111	75.14%	2,107	76.35%	2,107	79.85%
Denizli	1,930	71.97%	1,905	77.55%	1,899	80.43%	1,921	73.12%	1,926	60.07%	1,906	75.14%	1,898	76.35%	1,889	79.85%
Manisa	1,905	71.97%	1,878	77.55%	1,872	80.43%	1,896	73.12%	1,905	60.07%	1,880	75.14%	1,872	76.35%	1,862	79.85%
Aydın	1,855	71.97%	1,824	77.55%	1,815	80.43%	1,845	73.12%	1,863	60.07%	1,828	75.14%	1,818	76.35%	1,806	79.85%
Hatay	1,981	71.97%	1,960	77.55%	2,017	55.49%	1,973	73.12%	1,968	60.07%	1,959	75.14%	1,952	76.35%	2,005	62.30%
Eskişehir	2,020	71.97%	2,002	77.55%	2,048	55.49%	2,012	73.12%	2,001	60.07%	2,000	75.14%	1,994	76.35%	2,039	62.30%
Konya	2,220	71.97%	2,218	77.55%	2,202	55.49%	2,216	73.12%	2,168	60.07%	2,209	75.14%	2,240	26.61%	2,212	62.30%
Tekirdağ	2,084	71.97%	2,071	77.55%	2,097	55.49%	2,077	73.12%	2,054	60.07%	2,067	75.14%	2,189	26.61%	2,094	62.30%
Trabzon	2,175	71.97%	2,131	39.35%	2,167	55.49%	2,170	73.12%	2,130	60.07%	2,192	51.97%	2,223	26.61%	2,173	62.30%
Zonguldak	2,061	71.97%	2,069	39.35%	2,079	55.49%	2,054	73.12%	2,036	60.07%	2,110	51.97%	2,181	26.61%	2,075	62.30%
Sakarya	2,370	71.97%	2,238	39.35%	2,318	55.49%	2,368	73.12%	2,350	62.92%	2,333	51.97%	2,295	26.61%	2,342	62.30%
Erzurum	2,106	71.97%	2,093	39.35%	2,114	55.49%	2,100	73.12%	2,119	62.92%	2,142	51.97%	2,197	26.61%	2,113	62.30%
Diyarbakır	2,184	71.97%	2,136	39.35%	2,174	55.49%	2,179	73.12%	2,187	62.92%	2,199	51.97%	2,226	26.61%	2,173	56.03%
Ordu	2,403	71.97%	2,256	39.35%	2,343	55.49%	2,401	73.12%	2,378	62.92%	2,356	51.97%	2,307	26.61%	2,343	56.03%
Çanakkale	2,189	71.97%	2,139	39.35%	2,178	55.49%	2,037	8.74%	2,192	62.92%	2,202	51.97%	2,228	26.61%	2,177	56.03%
Kahramanmaraş	2,086	30.97%	2,037	39.35%	2,035	55.49%	2,014	8.74%	2,029	62.92%	2,068	51.97%	2,038	0.00%	2,032	56.03%
Şanlıurfa	2,180	30.97%	2,156	39.35%	2,203	55.49%	2,041	8.74%	2,219	62.92%	2,225	51.97%	2,038	0.00%	2,202	56.03%
Malatya	2,061	30.97%	2,005	39.35%	1,990	55.49%	2,007	8.74%	1,978	62.92%	2,026	51.97%	2,038	0.00%	1,987	56.03%
Edirne	2,060	30.97%	2,004	39.35%	1,987	55.49%	2,007	8.74%	1,975	62.92%	1,989	0.00%	2,038	0.00%	1,984	56.03%
Afyonkarahisar	2,124	30.97%	2,085	39.35%	2,102	55.49%	2,025	8.74%	2,105	62.92%	1,989	0.00%	2,038	0.00%	2,100	56.03%
Çorum	2,113	30.97%	2,071	39.35%	2,079	52.52%	2,022	8.74%	2,084	62.92%	1,989	0.00%	2,038	0.00%	1,992	0.00%
Elazığ	2,099	30.97%	2,054	39.35%	2,056	52.52%	2,018	8.74%	2,046	45.95%	1,989	0.00%	2,038	0.00%	1,992	0.00%
Giresun	2,097	30.97%	2,051	39.35%	2,051	52.52%	2,017	8.74%	2,042	45.95%	1,989	0.00%	2,038	0.00%	1,992	0.00%
Sivas	2,176	30.97%	2,151	39.35%	2,185	52.52%	2,039	8.74%	2,159	45.95%	1,989	0.00%	2,038	0.00%	1,992	0.00%
Batman	2,072	30.97%	2,019	39.35%	2,009	52.52%	2,010	8.74%	2,004	45.95%	1,989	0.00%	1,980	67.66%	1,992	0.00%
Kırklareli	2,031	30.97%	1,967	39.35%	1,939	52.52%	1,998	8.74%	1,943	45.95%	1,989	0.00%	1,890	67.66%	1,992	0.00%
Tokat	2,097	30.97%	2,050	39.35%	2,051	52.52%	2,017	8.74%	2,041	45.95%	1,989	0.00%	2,034	67.66%	1,992	0.00%
Uşak	2,087	30.97%	2,038	39.35%	2,034	52.52%	2,065	51.69%	2,027	45.95%	1,989	0.00%	2,013	67.66%	1,992	0.00%

Table D.2: Bühlmann premiums and credibility factors - 2

City	$k=3$		$k=4$		$k=5$		$k=6$		$k=7$		$k=8$		$k=9$		$k=10$	
Mardin	2,202	30.97%	2,354	46.21%	2,231	52.52%	2,259	51.69%	2,198	45.95%	2,282	54.05%	2,266	67.66%	2,298	61.70%
Kütahya	2,164	30.97%	2,296	46.21%	2,165	52.52%	2,194	51.69%	2,141	45.95%	2,214	54.05%	2,181	67.66%	2,220	61.70%
Van	2,278	30.97%	2,467	46.21%	2,359	52.52%	2,385	51.69%	2,311	45.95%	2,414	54.05%	2,431	67.66%	2,449	61.70%
Isparta	2,063	30.97%	2,146	46.21%	1,994	52.52%	2,025	51.69%	1,991	45.95%	2,038	54.05%	1,960	67.66%	2,019	61.70%
Düzce	2,231	30.97%	2,397	46.21%	2,279	52.52%	2,306	51.69%	2,338	17.93%	2,332	54.05%	2,328	67.66%	2,355	61.70%
Osmaniye	2,024	30.97%	2,088	46.21%	1,929	52.52%	1,961	51.69%	2,219	17.93%	1,971	54.05%	2,173	18.90%	1,943	61.70%
Adıyaman	2,140	30.97%	2,261	46.21%	2,125	52.52%	2,155	51.69%	2,286	17.93%	2,173	54.05%	2,244	18.90%	2,174	61.70%
Rize	2,276	30.97%	2,463	46.21%	2,355	52.52%	2,381	51.69%	2,364	17.93%	2,410	54.05%	2,327	18.90%	2,443	61.70%
Kastamonu	2,069	30.97%	2,155	46.21%	2,231	41.02%	2,036	51.69%	2,245	17.93%	2,050	54.05%	2,201	18.90%	2,238	18.04%
Amasya	2,219	30.97%	2,378	46.21%	2,429	41.02%	2,286	51.69%	2,331	17.93%	2,310	54.05%	2,292	18.90%	2,325	18.04%
Yalova	2,095	30.97%	2,194	46.21%	2,266	41.02%	2,080	51.69%	2,260	17.93%	2,297	42.20%	2,217	18.90%	2,253	18.04%
Aksaray	2,148	30.97%	2,272	46.21%	2,335	41.02%	2,167	51.69%	2,290	17.93%	2,368	42.20%	2,249	18.90%	2,284	18.04%
Nevşehir	2,169	30.97%	2,305	46.21%	2,364	41.02%	2,487	27.74%	2,303	17.93%	2,398	42.20%	2,262	18.90%	2,297	18.04%
Niğde	2,500	30.97%	2,798	46.21%	2,802	41.02%	2,783	27.74%	2,494	17.93%	2,848	42.20%	2,464	18.90%	2,489	18.04%
Bolu	2,550	35.41%	2,502	46.21%	2,539	41.02%	2,605	27.74%	2,379	17.93%	2,578	42.20%	2,593	58.44%	2,373	18.04%
Karabük	2,275	35.41%	2,142	46.21%	2,220	41.02%	2,389	27.74%	2,144	56.01%	2,249	42.20%	2,138	58.44%	2,233	18.04%
Siirt	2,349	35.41%	2,240	46.21%	2,306	41.02%	2,447	27.74%	2,263	56.01%	2,338	42.20%	2,262	58.44%	2,328	53.54%
Yozgat	3,028	35.41%	3,125	46.21%	3,092	41.02%	2,979	27.74%	3,335	56.01%	3,146	42.20%	3,381	58.44%	3,353	53.54%
Bilecik	2,437	35.41%	2,354	46.21%	2,408	41.02%	2,516	27.74%	2,402	56.01%	2,443	42.20%	2,407	58.44%	2,461	53.54%
Şırnak	2,527	35.41%	2,472	46.21%	2,512	41.02%	2,587	27.74%	2,544	56.01%	2,550	42.20%	2,555	58.44%	2,597	53.54%
Burdur	2,454	35.41%	2,444	25.49%	2,427	41.02%	2,529	27.74%	2,427	56.01%	2,377	10.47%	2,434	58.44%	2,485	53.54%
Ağrı	2,319	35.41%	2,347	25.49%	2,270	41.02%	2,423	27.74%	2,214	56.01%	2,337	10.47%	2,211	58.44%	2,281	53.54%
Artvin	2,546	35.41%	2,510	25.49%	2,533	41.02%	2,601	27.74%	2,573	56.01%	2,404	10.47%	2,585	58.44%	2,624	53.54%
Sinop	2,603	35.41%	2,551	25.49%	2,600	41.02%	2,646	27.74%	2,663	56.01%	2,421	10.47%	2,575	30.02%	2,711	53.54%
Erzincan	2,574	35.41%	2,531	25.49%	2,534	26.35%	2,623	27.74%	2,618	56.01%	2,412	10.47%	2,551	30.02%	2,529	32.25%
Kırşehir	2,286	35.41%	2,323	25.49%	2,320	26.35%	2,289	30.01%	2,162	56.01%	2,327	10.47%	2,307	30.02%	2,267	32.25%
Kars	2,606	35.41%	2,554	25.49%	2,558	26.35%	2,560	30.01%	2,573	26.36%	2,422	10.47%	2,578	30.02%	2,558	32.25%
Bitlis	2,509	35.41%	2,484	25.49%	2,486	26.35%	2,477	30.01%	2,501	26.36%	2,393	10.47%	2,496	30.02%	2,469	32.25%
Karaman	2,256	35.41%	2,302	25.49%	2,298	26.35%	2,263	30.01%	2,313	26.36%	2,318	10.47%	2,282	30.02%	2,239	32.25%
Muş	2,367	35.41%	2,381	25.49%	2,380	26.35%	2,357	30.01%	2,395	26.36%	2,351	10.47%	2,375	30.02%	2,340	32.25%
Bartın	2,754	35.41%	2,660	25.49%	2,668	26.35%	2,685	30.01%	2,683	26.36%	2,766	32.25%	2,703	30.02%	2,692	32.25%
Bingöl	2,436	35.41%	2,431	25.49%	2,432	26.35%	2,416	30.01%	2,447	26.36%	2,477	32.25%	2,434	30.02%	2,403	32.25%
Iğdır	2,440	35.41%	2,434	25.49%	2,435	26.35%	2,419	30.01%	2,450	26.36%	2,481	32.25%	2,449	29.87%	2,449	29.87%
Çankırı	2,331	35.41%	2,355	25.49%	2,353	26.35%	2,326	30.01%	2,368	26.36%	2,381	32.25%	2,357	29.87%	2,357	29.87%
Kırıkkale	2,516	35.41%	2,489	25.49%	2,491	26.35%	2,484	30.01%	2,506	26.36%	2,550	32.25%	2,513	29.87%	2,513	29.87%
Gümüşhane	2,594	35.41%	2,545	25.49%	2,550	26.35%	2,550	30.01%	2,565	26.36%	2,621	32.25%	2,579	29.87%	2,579	29.87%
Ardahan	2,698	35.41%	2,620	25.49%	2,627	26.35%	2,638	30.01%	2,642	26.36%	2,716	32.25%	2,667	29.87%	2,667	29.87%
Kilis	2,063	35.41%	2,163	25.49%	2,154	26.35%	2,099	30.01%	2,169	26.36%	2,137	32.25%	2,131	29.87%	2,131	29.87%
Bayburt	2,301	35.41%	2,334	25.49%	2,331	26.35%	2,301	30.01%	2,346	26.36%	2,354	32.25%	2,332	29.87%	2,332	29.87%
Hakkari	2,623	35.41%	2,566	25.49%	2,571	26.35%	2,574	30.01%	2,586	26.36%	2,647	32.25%	2,603	29.87%	2,603	29.87%
Tunceli	2,879	35.41%	2,750	25.49%	2,761	26.35%	2,791	30.01%	2,776	26.36%	2,880	32.25%	2,819	29.87%	2,819	29.87%

Table D.3: Bühlmann result comparison with target - 1 (difference)

	Number of Claim Grouping								Region	K-means			Hierarchical					
City	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$	$k=7$	$k=3$	$k=6$	$k=9$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
İstanbul	208	205	211	217	222	217	220	224	221	250	250	250	250	250	250	250	250	250
Ankara	213	203	205	217	214	208	209	216	334	165	188	188	284	185	185	185	188	188
İzmir	300	275	269	293	276	269	266	278	285	392	254	254	437	219	219	219	254	254
Bursa	213	193	190	210	197	191	189	199	219	256	192	132	327	351	343	332	332	351
Antalya	381	380	387	392	399	393	397	400	314	238	628	372	406	415	405	402	402	367
Adana	175	150	143	168	149	143	140	151	184	273	203	151	315	343	335	323	323	357
Kocaeli	83	92	104	100	117	112	118	118	101	259	223	185	67	69	58	58	58	3
Mersin	192	187	192	200	202	196	199	204	139	94	494	228	239	251	242	236	236	214
Gaziantep	245	215	206	234	212	205	200	197	246	385	312	266	405	436	429	416	416	462
Kayseri	48	34	35	49	43	37	25	23	181	28	31	162	132	151	142	134	134	134
Samsun	323	330	342	339	354	317	316	328	381	509	477	440	312	315	304	303	303	246
Balıkesir	298	296	303	308	246	288	285	290	312	172	122	306	332	342	332	328	328	298
Muğla	138	128	130	142	101	122	118	118	46	84	28	47	206	221	212	206	206	196
Denizli	368	343	337	359	364	344	336	327	352	458	389	336	504	531	524	512	512	544
Manisa	296	269	263	287	296	271	263	253	290	405	334	284	441	470	463	450	450	487
Aydın	339	308	299	329	347	312	302	290	352	484	411	365	502	533	526	512	512	560
Hatay	361	340	397	353	348	339	332	385	355	414	348	289	479	504	496	485	485	506
Eskişehir	181	163	209	173	162	161	155	200	332	205	142	79	286	308	300	290	290	303
Konya	235	233	217	231	183	224	255	227	326	507	508	247	270	281	271	267	267	237
Tekirdağ	389	376	402	382	359	372	494	399	398	366	307	238	472	490	481	473	473	473
Trabzon	139	95	131	134	94	156	187	137	225	49	452	184	190	203	193	188	188	168
Zonguldak	480	488	498	473	455	529	600	494	585	474	414	347	571	591	582	573	573	578
Sakarya	272	140	220	270	252	235	197	244	290	446	410	371	255	257	246	246	246	185
Erzurum	226	213	234	220	239	262	317	233	617	187	130	321	301	318	309	302	302	297
Diyarbakır	376	328	366	371	379	391	418	365	354	280	228	414	424	437	427	422	422	400
Ordu	497	350	437	495	472	450	401	437	547	650	605	376	468	469	457	457	457	390
Çanakkale	524	474	513	372	527	537	563	512	536	424	372	559	570	583	573	568	568	545
Kahramanmaraş	445	396	394	373	388	427	397	391	352	399	335	274	473	496	488	478	478	495
Şanlıurfa	434	410	457	295	473	479	292	456	450	352	747	486	510	521	511	506	506	477
Malatya	418	362	347	364	335	383	395	344	854	381	313	259	433	459	452	440	440	469
Edirne	442	386	369	389	357	371	420	366	328	405	337	283	456	483	475	463	463	493
Afyonkarahisar	145	184	167	244	164	280	231	169	256	206	264	335	98	80	89	97	97	99
Çorum	145	187	179	236	174	269	220	266	88	202	262	68	103	84	92	101	101	97
Elazığ	551	506	508	470	498	441	490	444	949	500	438	633	586	607	599	590	590	600
Giresun	621	575	575	541	566	513	562	516	662	570	508	444	654	676	667	658	658	669
Sivas	253	228	262	116	236	66	115	69	381	172	122	307	327	338	328	323	323	296
Batman	700	647	637	638	632	617	608	620	591	659	593	535	720	745	737	726	726	750
Kırklareli	450	386	358	417	362	408	309	411	295	424	352	305	449	479	472	459	459	503
Tokat	153	106	107	73	97	45	90	48	194	102	40	235	186	207	199	190	190	201
Uşak	319	270	266	297	259	221	245	224	192	272	208	147	347	370	362	352	352	368

Table D.4: Bühlmann result comparison with target - 2 (difference)

	Number of Claim Grouping								Region	K-means			Hierarchical					
City	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$	$k=7$	$k=3$	$k=6$	$k=9$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
Mardin	191	39	162	134	195	111	127	95	148	117	105	146	103	95	106	109	109	149
Kütahya	231	363	232	261	208	281	248	287	137	546	556	536	298	311	301	296	296	275
Van	123	312	204	230	156	259	276	294	342	417	361	314	250	247	236	237	237	159
Isparta	487	570	418	449	415	462	384	443	375	449	382	326	503	529	521	510	510	538
Düzce	96	70	48	21	11	5	1	28	73	207	178	142	6	10	1	2	2	57
Osmaniye	635	699	540	572	830	582	784	554	494	613	539	493	632	663	656	642	642	689
Adıyaman	403	524	388	418	549	436	507	437	374	336	280	206	459	475	465	459	459	449
Rize	328	515	407	433	416	462	379	495	539	622	567	685	453	451	439	440	440	364
Kastamonu	463	549	625	430	639	444	595	632	479	423	357	300	483	508	500	489	489	514
Amasya	506	665	716	573	618	597	579	612	663	409	789	544	602	607	597	595	595	546
Yalova	104	203	275	89	269	306	226	262	40	54	8	187	137	158	150	141	141	153
Aksaray	40	84	147	21	102	180	61	96	71	278	297	281	19	34	25	18	18	5
Nevşehir	466	602	661	784	600	695	559	594	591	780	787	523	537	549	539	534	534	510
Niğde	23	275	279	260	29	325	59	34	307	228	44	772	217	186	170	185	185	2
Bolu	806	758	795	861	635	834	849	629	793	847	777	565	696	690	678	681	681	592
Karabük	808	675	753	922	677	782	671	766	600	557	489	434	608	634	627	615	615	644
Siirt	801	692	758	899	715	790	714	780	531	516	458	387	626	644	635	627	627	624
Yozgat	1,131	1,228	1,195	1,082	1,438	1,249	1,484	1,456	1,287	1,035	720	1,398	1,174	1,114	1,094	1,122	1,122	1,719
Bilecik	821	738	792	900	786	827	791	845	673	895	882	853	674	682	672	669	669	628
Şırnak	1,046	991	1,031	1,106	1,063	1,069	1,074	1,116	934	1,094	1,035	815	928	925	914	915	915	836
Burdur	126	136	153	51	153	203	146	95	354	58	79	111	268	263	273	275	275	323
Ağrı	881	853	930	777	986	863	989	919	703	1,153	726	1,019	1,067	1,046	1,054	1,063	1,063	1,053
Artvin	35	71	48	20	8	177	4	43	52	7	61	112	148	152	165	162	162	249
Sinop	578	526	575	621	638	396	550	686	602	604	506	444	484	473	460	465	465	354
Erzincan	968	925	928	1,017	1,012	806	945	923	891	1,002	920	1,077	865	857	844	848	848	749
Kırşehir	561	598	595	564	437	602	582	542	409	305	238	438	365	390	382	371	371	396
Kars	767	715	719	721	734	583	739	719	658	364	340	1,456	674	662	649	655	655	542
Bitlis	348	323	325	316	340	232	335	308	336	401	351	308	224	223	212	212	212	141
Karaman	91	137	133	98	148	153	117	74	74	151	222	272	114	86	93	106	106	69
Muş	1,046	1,060	1,059	1,036	1,074	1,030	1,054	1,019	1,176	1,139	1,163	1,165	876	891	882	876	876	865
Bartın	567	473	481	498	496	579	516	505	697	549	375	668	522	493	477	491	491	1,429
Bingöl	160	155	156	140	171	201	158	127	221	233	222	193	12	20	10	7	7	33
Iğdır	2,733	2,739	2,738	2,754	2,723	2,692	2,724	2,724	2,676	2,661	2,675	2,925	2,880	2,872	2,883	2,886	2,886	2,927
Çankırı	479	503	501	474	516	529	505	505	345	202	141	75	297	317	308	300	300	305
Kırıkkale	903	876	878	871	893	937	900	900	847	954	901	677	781	780	768	770	770	695
Gümüşhane	742	693	698	698	713	769	727	727	760	771	678	617	646	635	622	628	628	520
Ardahan	418	340	347	358	362	436	387	387	217	28	271	1,015	355	333	318	329	329	176
Kilis	1,358	1,458	1,449	1,394	1,464	1,432	1,426	1,426	832	1,204	1,108	1,100	1,090	611	611	611	573	573
Bayburt	111	78	81	111	66	58	80	80	301	1	58	241	303	279	288	298	298	280
Hakkari	71	14	19	22	34	95	51	51	55	91	307	307	17	30	44	307	307	307
Tunceli	771	642	653	683	668	772	711	711	389	717	1,473	1,473	766	723	1,474	1,474	1,474	1,474

Table D.5: Bühlmann-Straub premiums and credibility factors - 1

City	$k=3$		$k=4$		$k=5$		$k=6$		$k=7$		$k=8$		$k=9$		$k=10$	
İstanbul	2,156	96.24%	2,154	95.41%	2,153	94.74%	2,152	92.79%	2,152	92.52%	2,150	92.00%	2,149	91.64%	2,148	89.59%
Ankara	2,021	90.20%	2,019	88.22%	2,018	86.63%	2,023	82.25%	2,022	81.66%	2,019	80.55%	2,019	79.80%	2,023	75.61%
İzmir	1,812	85.25%	1,815	82.46%	1,820	80.27%	1,840	74.41%	1,842	73.64%	1,841	72.21%	1,843	71.25%	1,862	66.05%
Bursa	1,912	75.48%	1,911	71.46%	1,915	68.43%	1,936	60.77%	1,936	59.82%	1,933	58.06%	1,933	56.91%	1,951	50.90%
Antalya	2,125	70.26%	2,111	65.78%	2,105	62.46%	2,105	54.32%	2,102	53.33%	2,092	51.52%	2,090	50.34%	2,091	44.31%
Adana	1,883	64.98%	1,884	60.15%	1,891	56.64%	1,922	48.28%	1,923	47.29%	1,919	45.49%	1,920	44.32%	1,944	38.45%
Kocaeli	2,202	57.20%	2,175	52.09%	2,163	48.48%	2,153	40.21%	2,148	39.26%	2,133	37.54%	2,128	36.44%	2,125	31.04%
Mersin	2,071	55.07%	2,055	49.92%	2,051	46.32%	2,060	38.15%	2,057	37.21%	2,047	35.53%	2,044	34.46%	2,053	29.21%
Gaziantep	1,837	52.03%	1,844	46.87%	1,855	43.30%	1,901	35.31%	1,902	34.40%	1,898	32.78%	1,900	31.75%	1,718	84.13%
Kayseri	2,024	48.11%	2,011	42.99%	2,010	39.49%	2,028	31.81%	2,026	30.95%	2,016	29.43%	2,021	78.49%	2,020	81.92%
Samsun	2,139	46.18%	2,112	41.11%	2,103	37.66%	2,103	30.16%	2,099	29.32%	2,222	79.20%	2,212	77.16%	2,220	80.75%
Balıkesir	2,056	45.51%	2,039	40.46%	2,036	37.03%	2,049	29.59%	2,060	69.88%	2,081	78.76%	2,075	76.68%	2,076	80.33%
Muğla	2,000	45.11%	1,988	40.07%	1,989	36.65%	2,012	29.26%	1,972	69.54%	1,983	78.49%	1,979	76.39%	1,976	80.07%
Denizli	1,927	45.03%	1,924	39.99%	1,931	36.58%	1,853	85.35%	1,861	69.47%	1,857	78.43%	1,857	76.33%	1,847	80.02%
Manisa	1,889	42.44%	1,890	37.49%	1,900	34.17%	1,770	83.98%	1,795	67.20%	1,782	76.59%	1,783	74.38%	1,770	78.28%
Aydın	1,879	41.85%	1,881	36.92%	1,892	33.63%	1,747	83.65%	1,777	66.66%	1,761	76.15%	1,764	73.91%	1,749	77.86%
Hatay	1,970	40.69%	1,962	35.82%	1,942	74.67%	1,927	82.99%	1,921	65.58%	1,925	75.28%	1,923	72.98%	1,943	78.00%
Eskişehir	1,958	40.43%	1,951	35.57%	1,920	74.46%	1,903	82.83%	1,902	65.34%	1,904	75.07%	1,902	72.77%	1,921	77.82%
Konya	2,089	39.37%	2,066	34.56%	2,163	73.61%	2,174	82.20%	2,114	64.33%	2,148	74.24%	2,178	67.64%	2,176	77.04%
Tekirdağ	1,976	38.99%	1,966	34.20%	1,951	73.30%	1,937	81.96%	1,928	63.96%	1,934	73.93%	1,983	67.29%	1,953	76.76%
Trabzon	2,047	38.50%	2,084	80.06%	2,086	72.89%	2,088	81.65%	2,045	63.49%	2,099	74.94%	2,108	66.83%	2,095	76.39%
Zonguldak	1,962	33.09%	1,896	76.04%	1,918	68.00%	1,896	77.86%	1,896	57.88%	1,928	70.26%	1,960	61.43%	1,919	71.88%
Sakarya	2,145	33.23%	2,316	76.15%	2,294	68.13%	2,326	77.96%	2,343	79.37%	2,316	70.39%	2,299	61.57%	2,316	72.01%
Erzurum	2,022	31.51%	2,034	74.69%	2,042	66.41%	2,038	76.59%	2,050	78.06%	2,056	68.73%	2,073	59.70%	2,050	70.40%
Diyarbakır	2,027	31.36%	2,047	74.56%	2,053	66.25%	2,051	76.46%	2,063	77.93%	2,068	68.57%	2,083	59.52%	2,038	57.79%
Ordu	2,072	26.55%	2,168	69.86%	2,160	60.83%	2,176	71.99%	2,194	73.64%	2,181	63.32%	2,181	53.78%	2,126	51.99%
Çanakkale	2,017	26.06%	2,025	69.33%	2,035	60.22%	1,975	41.86%	2,043	73.15%	2,052	62.73%	2,072	53.15%	2,020	51.36%
Kahramanmaraş	1,873	78.40%	1,884	68.67%	1,913	59.48%	1,889	41.12%	1,894	72.54%	1,925	62.01%	1,934	1.65%	1,916	50.60%
Şanlıurfa	2,099	76.96%	2,080	66.87%	2,084	57.48%	2,005	39.13%	2,103	70.87%	2,104	60.05%	1,938	1.52%	2,060	48.53%
Malatya	1,879	77.08%	1,890	67.01%	1,920	57.64%	1,893	39.29%	1,901	71.00%	1,932	60.20%	1,934	1.53%	1,921	48.69%
Edirne	1,884	76.91%	1,894	66.81%	1,924	57.41%	1,896	39.07%	1,906	70.81%	1,871	44.42%	1,934	1.51%	1,925	48.46%
Afyonkarahisar	1,929	76.22%	1,933	65.94%	1,958	56.46%	1,918	38.14%	1,948	70.00%	1,896	43.46%	1,935	1.46%	1,954	47.49%
Çorum	2,003	74.99%	1,996	64.43%	2,003	84.32%	1,954	36.59%	2,016	68.59%	1,936	41.84%	1,936	1.36%	1,951	54.12%
Elazığ	1,996	73.52%	1,990	62.65%	1,996	83.27%	1,949	34.82%	1,991	78.26%	1,930	39.98%	1,936	1.26%	1,944	52.20%
Giresun	1,913	73.06%	1,920	62.10%	1,902	82.95%	1,910	34.30%	1,902	77.87%	1,885	39.42%	1,935	1.24%	1,885	51.62%
Sivas	1,997	72.72%	1,991	61.69%	1,997	82.70%	1,949	33.91%	1,992	77.57%	1,930	39.01%	1,936	1.21%	1,944	51.19%
Batman	1,793	71.90%	1,819	60.72%	1,764	82.11%	1,855	33.00%	1,774	76.85%	1,821	38.04%	1,752	87.42%	1,801	50.17%
Kırklareli	1,798	70.57%	1,824	59.17%	1,768	81.13%	1,858	31.58%	1,778	75.67%	1,824	36.52%	1,756	86.69%	1,804	48.54%
Tokat	2,050	70.53%	2,035	59.12%	2,058	81.10%	1,971	31.54%	2,048	75.63%	1,955	36.48%	2,066	86.66%	1,978	48.49%
Uşak	1,799	69.29%	1,826	57.68%	1,768	80.18%	1,768	84.03%	1,778	74.53%	1,825	35.12%	1,755	85.97%	1,805	47.02%

Table D.6: Bühlmann-Straub premiums and credibility factors - 2

City	$k = 3$		$k = 4$		$k = 5$		$k = 6$		$k = 7$		$k = 8$		$k = 9$		$k = 10$	
Mardin	2,093	68.88%	2,140	69.14%	2,108	79.87%	2,125	83.77%	2,094	74.16%	2,129	82.49%	2,120	85.73%	2,133	85.47%
Kütahya	2,139	67.32%	2,188	67.58%	2,162	78.69%	2,183	82.77%	2,144	72.76%	2,187	81.43%	2,179	84.83%	2,193	84.56%
Van	2,352	67.49%	2,401	67.76%	2,411	78.82%	2,444	82.88%	2,374	72.92%	2,444	81.55%	2,447	84.93%	2,460	84.66%
Isparta	1,864	67.01%	1,912	67.28%	1,841	78.46%	1,845	82.57%	1,847	72.48%	1,854	81.22%	1,833	84.65%	1,847	84.38%
Düzce	2,251	66.82%	2,301	67.09%	2,294	78.31%	2,322	82.44%	2,299	71.30%	2,323	81.09%	2,322	84.54%	2,335	84.26%
Osmaniye	1,733	66.79%	1,781	67.06%	1,687	78.29%	1,683	82.42%	1,746	71.26%	1,695	81.07%	1,736	69.41%	1,682	84.24%
Adıyaman	1,844	66.02%	1,894	66.29%	1,817	77.70%	1,820	81.92%	1,866	70.56%	1,830	80.53%	1,853	68.68%	1,822	83.78%
Rize	2,337	65.46%	2,390	65.74%	2,399	77.26%	2,434	81.55%	2,394	70.04%	2,433	80.14%	2,366	68.14%	2,450	83.44%
Kastamonu	1,931	63.98%	1,984	64.26%	2,074	44.07%	1,927	80.55%	1,961	68.66%	1,936	79.09%	1,943	66.71%	2,079	9.50%
Amasya	2,100	64.09%	2,154	64.38%	2,190	44.19%	2,140	80.63%	2,142	68.77%	2,145	79.17%	2,120	66.83%	2,104	9.55%
Yalova	1,979	62.69%	2,035	62.98%	2,110	42.71%	1,988	79.67%	2,014	67.45%	2,140	26.71%	1,994	65.47%	2,087	9.04%
Aksaray	2,136	58.09%	2,199	58.39%	2,223	38.08%	2,195	76.37%	2,188	63.09%	2,210	23.11%	2,160	61.00%	2,109	7.57%
Nevşehir	2,107	58.35%	2,169	58.65%	2,204	38.33%	2,265	23.43%	2,156	63.34%	2,198	23.30%	2,130	61.25%	2,105	7.65%
Niğde	2,068	57.83%	2,132	58.13%	2,179	37.82%	2,250	23.05%	2,115	62.84%	2,184	22.92%	2,090	60.74%	2,100	7.50%
Bolu	2,378	67.61%	2,304	57.56%	2,291	37.28%	2,319	22.64%	2,301	62.30%	2,251	22.51%	2,358	46.75%	2,122	7.34%
Karabük	2,003	65.11%	1,984	54.81%	2,088	34.70%	2,198	20.74%	2,068	49.81%	2,129	20.62%	2,104	43.98%	2,082	6.62%
Siirt	2,139	62.33%	2,095	51.81%	2,161	32.03%	2,244	18.83%	2,173	46.80%	2,173	18.72%	2,199	41.03%	2,287	25.42%
Yozgat	2,635	61.29%	2,505	50.71%	2,413	31.08%	2,392	18.17%	2,545	45.71%	2,320	18.06%	2,524	39.97%	2,488	24.59%
Bilecik	2,330	61.42%	2,252	50.85%	2,258	31.19%	2,301	18.25%	2,316	45.84%	2,230	18.14%	2,324	40.10%	2,365	24.69%
Şırnak	2,194	60.11%	2,139	49.47%	2,190	30.03%	2,261	17.44%	2,216	44.48%	2,190	17.34%	2,237	38.79%	2,312	23.68%
Burdur	2,400	60.11%	2,426	73.48%	2,292	30.03%	2,321	17.44%	2,368	44.48%	2,372	53.11%	2,370	38.79%	2,393	23.68%
Ağrı	2,138	58.96%	2,106	72.53%	2,162	29.03%	2,246	16.76%	2,175	43.31%	2,141	51.92%	2,202	37.66%	2,292	22.83%
Artvin	2,494	57.74%	2,545	71.52%	2,335	28.01%	2,346	16.07%	2,436	42.08%	2,454	50.66%	2,428	36.49%	2,429	21.96%
Sinop	2,378	56.82%	2,402	70.75%	2,278	27.26%	2,313	15.57%	2,351	41.17%	2,351	49.73%	2,393	64.16%	2,385	21.32%
Erzincan	2,350	56.57%	2,367	70.54%	2,365	75.04%	2,306	15.44%	2,330	40.92%	2,327	49.47%	2,361	63.93%	2,357	66.79%
Kırşehir	2,116	55.95%	2,074	70.02%	2,053	74.57%	2,041	77.82%	2,161	40.32%	2,122	48.84%	2,096	63.34%	2,080	66.23%
Kars	2,136	53.16%	2,097	67.60%	2,075	72.37%	2,063	75.82%	2,069	75.97%	2,140	46.04%	2,118	60.69%	2,102	63.67%
Bitlis	2,366	51.42%	2,391	66.05%	2,389	70.95%	2,393	74.52%	2,399	74.68%	2,338	44.31%	2,381	59.01%	2,378	62.04%
Karaman	2,067	50.43%	2,004	65.16%	1,974	70.13%	1,956	73.76%	1,962	73.92%	2,081	43.34%	2,037	58.05%	2,016	61.10%
Muş	2,291	50.57%	2,294	65.28%	2,286	70.24%	2,284	73.86%	2,291	74.02%	2,273	43.47%	2,295	58.18%	2,287	61.23%
Bartın	2,626	49.60%	2,729	64.41%	2,754	69.43%	2,778	73.11%	2,785	73.28%	2,849	80.76%	2,682	57.24%	2,695	60.31%
Bingöl	2,336	48.76%	2,353	63.63%	2,348	68.72%	2,350	72.45%	2,357	72.61%	2,378	80.23%	2,347	56.42%	2,342	59.50%
Iğdır	2,276	47.52%	2,275	62.47%	2,264	67.63%	2,261	71.44%	2,268	71.61%	2,279	79.43%	2,262	80.33%	2,262	80.33%
Çankırı	2,102	45.46%	2,042	60.51%	2,010	65.79%	1,992	69.72%	1,999	69.90%	1,980	78.04%	1,957	78.99%	1,957	78.99%
Kırıkkale	2,305	42.48%	2,313	57.58%	2,304	63.02%	2,304	67.10%	2,312	67.29%	2,331	75.90%	2,311	76.90%	2,311	76.90%
Gümüşhane	2,386	36.14%	2,429	50.99%	2,431	56.64%	2,440	60.99%	2,450	61.19%	2,495	70.71%	2,473	71.85%	2,473	71.85%
Ardahan	2,233	33.24%	2,212	47.79%	2,189	53.47%	2,179	57.90%	2,188	58.11%	2,193	67.98%	2,164	69.19%	2,164	69.19%
Kilis	2,006	25.73%	1,865	38.90%	1,788	44.43%	1,735	48.90%	1,745	49.11%	1,656	59.63%	1,607	60.97%	1,607	60.97%
Bayburt	2,240	24.25%	2,218	37.05%	2,191	42.49%	2,178	46.93%	2,190	47.14%	2,197	57.72%	2,158	59.08%	2,158	59.08%
Hakkari	2,546	24.07%	2,687	36.82%	2,728	42.25%	2,771	46.69%	2,786	46.90%	2,928	57.48%	2,906	58.84%	2,906	58.84%
Tunceli	2,409	15.43%	2,489	25.12%	2,504	29.63%	2,529	33.51%	2,546	33.70%	2,662	43.76%	2,625	45.14%	2,625	45.14%

Table D.7: Bühlmann-Straub result comparison with target - 1 (difference)

	Number of Claim Grouping								Region	K-means			Hierarchical					
City	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$	$k=7$	$k=3$	$k=6$	$k=9$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
İstanbul	140	138	137	136	136	134	133	132	136	146	146	146	146	146	146	146	146	146
Ankara	115	113	112	117	116	113	113	117	131	111	97	97	117	101	101	101	97	97
İzmir	184	187	192	212	214	213	215	234	155	162	182	182	158	173	173	173	182	182
Bursa	131	130	134	155	155	152	152	170	237	106	101	56	109	107	107	107	107	107
Antalya	258	244	238	238	235	225	223	224	289	262	352	253	299	302	302	302	302	301
Adana	138	139	146	177	178	174	175	199	87	95	89	85	92	87	87	87	87	86
Kocaeli	82	109	121	131	136	151	156	159	134	62	10	74	26	38	37	36	36	34
Mersin	76	60	56	65	62	52	49	58	102	68	220	125	117	122	121	121	121	119
Gaziantep	220	227	238	284	285	281	283	101	94	143	141	208	117	104	104	105	105	104
Kayseri	8	21	22	4	6	16	11	12	5	33	55	88	11	12	12	12	12	10
Samsun	110	83	74	74	70	193	183	191	250	309	242	329	218	232	231	230	230	227
Balıkesir	129	112	109	122	133	154	148	149	174	113	86	193	175	180	180	179	179	177
Muğla	11	1	0	23	17	6	10	13	63	25	46	152	11	10	10	10	10	8
Denizli	365	362	369	291	299	295	295	285	268	304	292	270	307	299	299	299	299	298
Manisa	280	281	291	161	186	173	174	161	163	201	193	220	183	169	169	169	169	168
Aydın	363	365	376	231	261	245	248	233	240	279	272	312	255	239	239	240	240	239
Hatay	350	342	322	307	301	305	303	323	306	299	280	214	323	318	318	318	318	316
Eskişehir	119	112	81	64	63	65	63	82	198	63	46	6	81	75	75	75	75	73
Konya	104	81	178	189	129	163	193	191	52	347	268	135	192	204	203	202	202	199
Tekirdağ	281	271	256	242	233	239	288	258	385	229	210	139	257	253	253	253	253	251
Trabzon	11	48	50	52	9	63	72	59	88	11	196	84	59	66	65	64	64	62
Zonguldak	381	315	337	315	315	347	379	338	352	317	297	251	337	329	328	328	328	326
Sakarya	47	218	196	228	245	218	201	218	29	253	193	260	221	249	248	246	246	241
Erzurum	142	154	162	158	170	176	193	170	190	104	74	240	170	174	174	173	173	170
Diyarbakır	219	239	245	243	255	260	275	230	221	183	153	312	255	260	259	258	258	255
Ordu	166	262	254	270	288	275	275	220	329	434	356	214	273	293	292	290	290	285
Çanakkale	352	360	370	310	378	387	407	355	431	307	276	455	380	384	384	383	383	379
Kahramanmaraş	232	243	272	248	253	284	293	275	247	250	229	191	271	260	260	260	260	257
Şanlıurfa	353	334	338	259	357	358	192	314	310	255	502	374	352	364	363	361	361	357
Malatya	236	247	277	250	258	289	291	278	257	251	229	189	276	265	265	265	265	261
Edirne	266	276	306	278	288	253	316	307	467	279	256	214	306	295	295	294	294	291
Afyonkarahisar	340	336	311	351	321	373	334	315	410	351	376	436	308	314	314	315	315	319
Çorum	255	262	255	304	242	322	322	307	196	306	337	138	237	235	236	237	237	241
Elazığ	448	442	448	401	443	382	388	396	487	400	369	572	469	470	469	468	468	464
Giresun	437	444	426	434	426	409	459	409	507	434	409	356	476	467	467	466	466	462
Sivas	74	68	74	26	69	7	13	21	114	25	7	197	95	97	96	95	95	90
Batman	421	447	392	483	402	449	380	429	425	486	467	458	487	464	464	465	465	461
Kırklareli	217	243	187	277	197	243	175	223	502	281	261	250	284	261	261	262	262	258
Tokat	106	91	114	27	104	11	122	34	171	24	11	176	117	125	124	123	123	117
Uşak	31	58	0	0	10	57	13	37	30	95	75	63	100	77	77	77	77	74

Table D.8: Bühlmann-Straub result comparison with target - 2 (difference)

	Number of Claim Grouping								Region	K-means			Hierarchical					
City	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$	$k=7$	$k=3$	$k=6$	$k=9$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$
Mardin	300	253	285	268	299	264	273	260	353	55	139	273	299	285	286	288	288	294
Kütahya	206	255	229	250	211	254	246	260	3	408	327	425	195	216	215	213	213	206
Van	197	246	256	289	219	289	292	305	305	196	129	203	135	184	182	178	178	169
Isparta	288	336	265	269	271	278	257	271	312	315	290	255	345	330	329	329	329	325
Düzce	76	26	33	5	28	4	5	8	21	19	55	31	114	78	80	83	83	91
Osmaniye	344	392	298	294	357	306	347	293	411	450	433	441	434	401	401	402	402	399
Adıyaman	107	157	80	83	129	93	116	85	101	146	123	94	171	152	152	152	152	148
Rize	389	442	451	486	446	485	418	502	439	403	334	576	328	377	376	371	371	362
Kastamonu	325	378	468	321	355	330	337	473	414	311	281	226	369	362	362	361	361	355
Amasya	387	441	477	427	429	432	407	391	461	268	544	407	387	404	403	400	400	393
Yalova	12	44	119	3	23	149	3	96	104	57	89	129	20	20	20	18	18	12
Aksaray	52	11	35	7	0	22	28	79	151	155	74	170	64	39	40	43	43	51
Nevşehir	404	466	501	562	453	495	427	402	334	639	557	417	401	421	420	417	417	409
Niğde	455	391	344	273	408	339	433	423	486	560	267	217	446	431	433	435	435	443
Bolu	634	560	547	575	557	507	614	378	566	603	528	376	453	494	493	488	488	479
Karabük	536	517	621	731	601	662	637	615	561	444	413	365	505	497	496	495	495	489
Siirt	591	547	613	696	625	625	651	739	418	396	359	284	499	508	507	505	505	497
Yozgat	738	608	516	495	648	423	627	591	140	457	390	422	423	501	498	491	491	999
Bilecik	714	636	642	685	700	614	708	749	488	730	651	742	535	571	569	565	565	555
Şırnak	713	658	709	780	735	709	756	831	512	861	778	639	593	610	608	606	606	597
Burdur	180	154	288	259	212	208	210	187	475	233	308	222	394	349	351	355	355	366
Ağrı	1,062	1,094	1,038	954	1,025	1,059	998	908	1,092	1,261	944	1,080	1,158	1,149	1,150	1,153	1,153	1,161
Artvin	87	36	246	235	145	127	153	152	204	231	304	223	351	293	295	301	301	313
Sinop	353	377	253	288	326	326	368	360	270	322	245	333	141	183	181	176	176	165
Erzincan	744	761	759	700	724	721	755	751	724	740	662	896	544	582	581	576	576	565
Kırşehir	391	349	328	316	436	397	371	355	312	206	170	395	302	307	306	304	304	295
Kars	297	258	236	224	230	301	279	263	262	94	57	468	195	202	201	199	199	190
Bitlis	205	230	228	232	238	177	220	217	183	185	107	197	15	25	23	19	19	7
Karaman	98	161	191	209	203	84	128	149	128	249	283	333	168	171	172	174	174	182
Muş	970	973	965	963	970	952	974	966	943	1024	944	1170	787	816	814	810	810	799
Bartın	439	542	567	591	598	662	495	508	263	165	94	407	82	159	156	149	149	709
Bingöl	60	77	72	74	81	102	71	66	35	70	9	82	150	115	117	121	121	133
Iğdır	2,897	2,898	2,909	2,912	2,905	2,894	2,911	2,911	2,927	2,828	2,909	3,053	3,078	3,052	3,053	3,057	3,057	3,068
Çankırı	250	190	158	140	147	128	105	105	185	68	33	20	158	159	158	155	155	146
Kırıkkale	692	700	691	691	699	718	698	698	424	732	652	507	487	517	515	511	511	499
Gümüşhane	534	577	579	588	598	643	621	621	423	495	416	506	273	313	311	306	306	292
Ardahan	47	68	91	101	92	87	116	116	90	345	18	29	223	207	208	212	212	224
Kilis	1,301	1,160	1,083	1,030	1,040	951	902	902	1,131	1,191	1,159	1,126	1,255	1,182	1,182	1,182	476	476
Bayburt	172	194	221	234	222	215	254	254	232	67	150	292	362	347	348	352	352	364
Hakkari	6	135	176	219	234	376	354	354	23	203	796	796	384	324	327	796	796	796
Tunceli	301	381	396	421	438	554	517	517	266	239	943	943	5	33	943	943	943	943

Table D.9: Bühlmann premiums and credibility factors - 1

	Region		K-means						Hierarchical											
City	$k=7$		$k=3$		$k=6$		$k=9$		$k=2$		$k=3$		$k=4$		$k=5$		$k=6$		$k=7$	
İstanbul	2,237	74.25%	2,266	0.00%	2,266	0.00%	2,266	0.00%	2,266	0.00%	2,266	0.00%	2,266	0.00%	2,266	0.00%	2,266	0.00%	2,266	0.00%
Ankara	2,240	50.18%	2,071	19.11%	2,094	80.07%	2,094	80.07%	2,190	46.84%	2,091	91.96%	2,091	91.96%	2,091	91.96%	2,094	80.07%	2,094	80.07%
İzmir	1,913	44.13%	2,020	19.11%	1,882	80.07%	1,882	80.07%	2,065	46.84%	1,847	91.96%	1,847	91.96%	1,847	91.96%	1,882	80.07%	1,882	80.07%
Bursa	2,000	74.25%	2,037	19.11%	1,973	23.87%	1,913	16.08%	2,108	46.84%	2,132	91.96%	2,124	91.96%	2,113	91.96%	2,113	44.05%	2,132	28.85%
Antalya	2,181	55.24%	2,105	19.11%	2,495	7.16%	2,239	19.61%	2,273	46.84%	2,282	42.78%	2,272	42.19%	2,269	44.05%	2,269	44.05%	2,234	28.85%
Adana	1,929	55.24%	2,018	19.11%	1,948	23.87%	1,896	16.08%	2,060	46.84%	2,088	42.78%	2,080	42.19%	2,068	44.05%	2,068	44.05%	2,102	28.85%
Kocaeli	2,385	74.25%	2,543	25.27%	2,507	7.16%	2,469	0.00%	2,351	46.84%	2,353	42.78%	2,342	42.19%	2,342	44.05%	2,342	44.05%	2,281	28.85%
Mersin	2,134	55.24%	2,089	19.11%	2,489	7.16%	2,223	19.61%	2,234	46.84%	2,246	42.78%	2,237	42.19%	2,231	44.05%	2,231	44.05%	2,209	28.85%
Gaziantep	1,863	66.93%	2,002	19.11%	1,929	23.87%	1,883	16.08%	2,022	46.84%	2,053	42.78%	2,046	42.19%	2,033	44.05%	2,033	44.05%	2,079	28.85%
Kayseri	2,213	50.18%	2,060	19.11%	2,001	23.87%	2,194	19.61%	2,164	46.84%	2,183	42.78%	2,174	42.19%	2,166	44.05%	2,166	44.05%	2,166	28.85%
Samsun	2,410	60.40%	2,538	25.27%	2,506	7.16%	2,469	0.00%	2,341	46.84%	2,344	42.78%	2,333	42.19%	2,332	44.05%	2,332	44.05%	2,275	28.85%
Balıkesir	2,239	74.25%	2,099	19.11%	2,049	23.87%	2,233	19.61%	2,259	46.84%	2,269	42.78%	2,259	42.19%	2,255	44.05%	2,255	44.05%	2,225	28.85%
Muğla	2,035	44.13%	2,073	19.11%	2,017	23.87%	1,942	16.08%	2,195	46.84%	2,210	42.78%	2,201	42.19%	2,195	44.05%	2,195	44.05%	2,185	28.85%
Denizli	1,914	44.13%	2,020	19.11%	1,951	23.87%	1,898	16.08%	2,066	46.84%	2,093	42.78%	2,086	42.19%	2,074	44.05%	2,074	44.05%	2,106	28.85%
Manisa	1,899	44.13%	2,014	19.11%	1,943	23.87%	1,893	16.08%	2,050	46.84%	2,079	42.78%	2,072	42.19%	2,059	44.05%	2,059	44.05%	2,096	28.85%
Aydın	1,868	44.13%	2,000	19.11%	1,927	23.87%	1,881	16.08%	2,018	46.84%	2,049	42.78%	2,042	42.19%	2,028	44.05%	2,028	44.05%	2,076	28.85%
Hatay	1,975	55.24%	2,034	19.11%	1,968	23.87%	1,909	16.08%	2,099	46.84%	2,124	42.78%	2,116	42.19%	2,105	44.05%	2,105	44.05%	2,126	28.85%
Eskişehir	2,171	50.18%	2,044	19.11%	1,981	23.87%	1,918	16.08%	2,125	46.84%	2,147	42.78%	2,139	42.19%	2,129	44.05%	2,129	44.05%	2,142	28.85%
Konya	2,311	50.18%	2,492	25.27%	2,493	7.16%	2,232	19.61%	2,255	46.84%	2,266	42.78%	2,256	42.19%	2,252	44.05%	2,252	44.05%	2,222	28.85%
Tekirdağ	2,093	74.25%	2,061	19.11%	2,002	23.87%	1,933	16.08%	2,167	46.84%	2,185	42.78%	2,176	42.19%	2,168	44.05%	2,168	44.05%	2,168	28.85%
Trabzon	2,261	60.40%	2,085	19.11%	2,488	7.16%	2,220	19.61%	2,226	46.84%	2,239	42.78%	2,229	42.19%	2,224	44.05%	2,224	44.05%	2,204	28.85%
Zonguldak	2,166	60.40%	2,055	19.11%	1,995	23.87%	1,928	16.08%	2,152	46.84%	2,172	42.78%	2,163	42.19%	2,154	44.05%	2,154	44.05%	2,159	28.85%
Sakarya	2,388	74.25%	2,544	25.27%	2,508	7.16%	2,469	0.00%	2,353	46.84%	2,355	42.78%	2,344	42.19%	2,344	44.05%	2,344	44.05%	2,283	28.85%
Erzurum	2,497	0.00%	2,067	19.11%	2,010	23.87%	2,201	19.61%	2,181	46.84%	2,198	42.78%	2,189	42.19%	2,182	44.05%	2,182	44.05%	2,177	28.85%
Diyarbakır	2,162	66.93%	2,088	19.11%	2,036	23.87%	2,222	19.61%	2,232	46.84%	2,245	42.78%	2,235	42.19%	2,230	44.05%	2,230	44.05%	2,208	28.85%
Ordu	2,453	60.40%	2,556	25.27%	2,511	7.16%	2,282	19.61%	2,374	46.84%	2,375	42.78%	2,363	42.19%	2,363	44.05%	2,363	44.05%	2,296	28.85%
Çanakkale	2,201	74.25%	2,089	19.11%	2,037	23.87%	2,224	19.61%	2,235	46.84%	2,248	42.78%	2,238	42.19%	2,233	44.05%	2,233	44.05%	2,210	28.85%
Kahramanmaraş	1,993	55.24%	2,040	19.11%	1,976	23.87%	1,915	16.08%	2,114	46.84%	2,137	42.78%	2,129	42.19%	2,119	44.05%	2,119	44.05%	2,136	28.85%
Şanlıurfa	2,196	66.93%	2,098	19.11%	2,493	7.16%	2,232	19.61%	2,256	46.84%	2,267	42.78%	2,257	42.19%	2,252	44.05%	2,252	44.05%	2,223	28.85%
Malatya	2,497	0.00%	2,024	19.11%	1,956	23.87%	1,902	16.08%	2,076	46.84%	2,102	42.78%	2,095	42.19%	2,083	44.05%	2,083	44.05%	2,112	28.85%
Edirne	1,946	74.25%	2,023	19.11%	1,955	23.87%	1,901	16.08%	2,074	46.84%	2,101	42.78%	2,093	42.19%	2,081	44.05%	2,081	44.05%	2,111	28.85%
Afyonkarahisar	2,013	44.13%	2,063	19.11%	2,005	23.87%	1,934	16.08%	2,171	46.84%	2,189	42.78%	2,180	42.19%	2,172	44.05%	2,172	44.05%	2,170	28.85%
Çorum	2,170	60.40%	2,056	19.11%	1,996	23.87%	2,190	19.61%	2,155	46.84%	2,174	42.78%	2,166	42.19%	2,157	44.05%	2,157	44.05%	2,161	28.85%
Elazığ	2,497	0.00%	2,048	19.11%	1,986	23.87%	2,181	19.61%	2,134	46.84%	2,155	42.78%	2,147	42.19%	2,138	44.05%	2,138	44.05%	2,148	28.85%
Giresun	2,138	60.40%	2,046	19.11%	1,984	23.87%	1,920	16.08%	2,130	46.84%	2,152	42.78%	2,143	42.19%	2,134	44.05%	2,134	44.05%	2,145	28.85%
Sivas	2,304	50.18%	2,095	19.11%	2,045	23.87%	2,230	19.61%	2,250	46.84%	2,261	42.78%	2,251	42.19%	2,246	44.05%	2,246	44.05%	2,219	28.85%
Batman	1,963	66.93%	2,031	19.11%	1,965	23.87%	1,907	16.08%	2,092	46.84%	2,117	42.78%	2,109	42.19%	2,098	44.05%	2,098	44.05%	2,122	28.85%
Kırklareli	1,876	74.25%	2,005	19.11%	1,933	23.87%	1,886	16.08%	2,030	46.84%	2,060	42.78%	2,053	42.19%	2,040	44.05%	2,040	44.05%	2,084	28.85%
Tokat	2,138	60.40%	2,046	19.11%	1,984	23.87%	2,179	19.61%	2,130	46.84%	2,151	42.78%	2,143	42.19%	2,134	44.05%	2,134	44.05%	2,145	28.85%
Uşak	1,960	44.13%	2,040	19.11%	1,976	23.87%	1,915	16.08%	2,115	46.84%	2,138	42.78%	2,130	42.19%	2,120	44.05%	2,120	44.05%	2,136	28.85%

Table D.10: Bühlmann premiums and credibility factors - 2

	Region		K-means						Hierarchical											
City	$k=7$		$k=3$		$k=6$		$k=9$		$k=2$		$k=3$		$k=4$		$k=5$		$k=6$		$k=7$	
Mardin	2,245	66.93%	2,510	25.27%	2,498	7.16%	2,247	19.61%	2,290	46.84%	2,298	42.78%	2,287	42.19%	2,284	44.05%	2,284	44.05%	2,244	28.85%
Kütahya	2,070	44.13%	2,479	25.27%	2,489	7.16%	2,469	0.00%	2,231	46.84%	2,244	42.78%	2,234	42.19%	2,229	44.05%	2,229	44.05%	2,208	28.85%
Van	2,497	0.00%	2,572	25.27%	2,516	7.16%	2,469	0.00%	2,405	46.84%	2,402	42.78%	2,391	42.19%	2,392	44.05%	2,392	44.05%	2,314	28.85%
Isparta	1,951	55.24%	2,025	19.11%	1,958	23.87%	1,902	16.08%	2,079	46.84%	2,105	42.78%	2,097	42.19%	2,086	44.05%	2,086	44.05%	2,114	28.85%
Düzce	2,400	60.40%	2,534	25.27%	2,505	7.16%	2,469	0.00%	2,333	46.84%	2,337	42.78%	2,326	42.19%	2,325	44.05%	2,325	44.05%	2,270	28.85%
Osmaniye	1,883	55.24%	2,002	19.11%	1,928	23.87%	1,882	16.08%	2,021	46.84%	2,052	42.78%	2,045	42.19%	2,031	44.05%	2,031	44.05%	2,078	28.85%
Adıyaman	2,111	66.93%	2,073	19.11%	2,017	23.87%	1,943	16.08%	2,196	46.84%	2,212	42.78%	2,202	42.19%	2,196	44.05%	2,196	44.05%	2,186	28.85%
Rize	2,487	60.40%	2,570	25.27%	2,515	7.16%	2,633	33.77%	2,401	46.84%	2,399	42.78%	2,387	42.19%	2,388	44.05%	2,388	44.05%	2,312	28.85%
Kastamonu	2,085	60.40%	2,029	19.11%	1,963	23.87%	1,906	16.08%	2,089	46.84%	2,114	42.78%	2,106	42.19%	2,095	44.05%	2,095	44.05%	2,120	28.85%
Amasya	2,376	60.40%	2,122	19.11%	2,502	7.16%	2,257	19.61%	2,315	46.84%	2,320	42.78%	2,310	42.19%	2,308	44.05%	2,308	44.05%	2,259	28.85%
Yalova	2,031	74.25%	2,045	19.11%	1,983	23.87%	2,178	19.61%	2,128	46.84%	2,149	42.78%	2,141	42.19%	2,132	44.05%	2,132	44.05%	2,144	28.85%
Aksaray	2,259	50.18%	2,466	25.27%	2,485	7.16%	2,469	0.00%	2,207	46.84%	2,222	42.78%	2,213	42.19%	2,206	44.05%	2,206	44.05%	2,193	28.85%
Nevşehir	2,294	50.18%	2,483	25.27%	2,490	7.16%	2,226	19.61%	2,240	46.84%	2,252	42.78%	2,242	42.19%	2,237	44.05%	2,237	44.05%	2,213	28.85%
Niğde	2,830	50.18%	2,295	19.11%	2,567	7.16%	3,295	0.00%	2,740	46.84%	2,709	42.78%	2,693	42.19%	2,708	44.05%	2,708	44.05%	2,521	28.85%
Bolu	2,537	60.40%	2,591	25.27%	2,521	7.16%	2,309	19.61%	2,440	46.84%	2,434	42.78%	2,422	42.19%	2,425	44.05%	2,425	44.05%	2,336	28.85%
Karabük	2,067	60.40%	2,024	19.11%	1,956	23.87%	1,901	16.08%	2,075	46.84%	2,101	42.78%	2,094	42.19%	2,082	44.05%	2,082	44.05%	2,111	28.85%
Siirt	2,079	66.93%	2,064	19.11%	2,006	23.87%	1,935	16.08%	2,174	46.84%	2,192	42.78%	2,183	42.19%	2,175	44.05%	2,175	44.05%	2,172	28.85%
Yozgat	3,184	50.18%	2,932	25.27%	2,617	7.16%	3,295	0.00%	3,071	46.84%	3,011	42.78%	2,991	42.19%	3,019	44.05%	3,019	44.05%	3,616	28.85%
Bilecik	2,289	74.25%	2,511	25.27%	2,498	7.16%	2,469	0.00%	2,290	46.84%	2,298	42.78%	2,288	42.19%	2,285	44.05%	2,285	44.05%	2,244	28.85%
Şırnak	2,415	66.93%	2,575	25.27%	2,516	7.16%	2,296	19.61%	2,409	46.84%	2,406	42.78%	2,395	42.19%	2,396	44.05%	2,396	44.05%	2,317	28.85%
Burdur	2,226	55.24%	2,522	25.27%	2,501	7.16%	2,469	0.00%	2,312	46.84%	2,317	42.78%	2,307	42.19%	2,305	44.05%	2,305	44.05%	2,257	28.85%
Ağrı	2,497	0.00%	2,047	19.11%	2,474	7.16%	2,181	19.61%	2,133	46.84%	2,154	42.78%	2,146	42.19%	2,137	44.05%	2,137	44.05%	2,147	28.85%
Artvin	2,529	60.40%	2,588	25.27%	2,520	7.16%	2,469	0.00%	2,433	46.84%	2,429	42.78%	2,416	42.19%	2,419	44.05%	2,419	44.05%	2,332	28.85%
Sinop	2,627	60.40%	2,629	25.27%	2,531	7.16%	2,469	0.00%	2,509	46.84%	2,498	42.78%	2,485	42.19%	2,490	44.05%	2,490	44.05%	2,379	28.85%
Erzincan	2,497	0.00%	2,608	25.27%	2,526	7.16%	2,683	33.77%	2,471	46.84%	2,463	42.78%	2,450	42.19%	2,454	44.05%	2,454	44.05%	2,355	28.85%
Kırşehir	2,134	50.18%	2,030	19.11%	1,963	23.87%	2,163	19.61%	2,090	46.84%	2,115	42.78%	2,107	42.19%	2,096	44.05%	2,096	44.05%	2,121	28.85%
Kars	2,497	0.00%	2,203	19.11%	2,179	23.87%	3,295	0.00%	2,513	46.84%	2,501	42.78%	2,488	42.19%	2,494	44.05%	2,494	44.05%	2,381	28.85%
Bitlis	2,497	0.00%	2,562	25.27%	2,512	7.16%	2,469	0.00%	2,385	46.84%	2,384	42.78%	2,373	42.19%	2,373	44.05%	2,373	44.05%	2,302	28.85%
Karaman	2,091	50.18%	2,014	19.11%	1,943	23.87%	1,893	16.08%	2,051	46.84%	2,079	42.78%	2,072	42.19%	2,059	44.05%	2,059	44.05%	2,096	28.85%
Muş	2,497	0.00%	2,460	25.27%	2,484	7.16%	2,486	33.77%	2,197	46.84%	2,212	42.78%	2,203	42.19%	2,197	44.05%	2,197	44.05%	2,186	28.85%
Bartın	2,884	60.40%	2,736	25.27%	2,562	7.16%	2,855	33.77%	2,709	46.84%	2,680	42.78%	2,664	42.19%	2,678	44.05%	2,678	44.05%	3,616	28.85%
Bingöl	2,497	0.00%	2,509	25.27%	2,498	7.16%	2,469	0.00%	2,288	46.84%	2,296	42.78%	2,286	42.19%	2,283	44.05%	2,283	44.05%	2,243	28.85%
Iğdır	2,497	0.00%	2,512	25.27%	2,498	7.16%	2,248	19.61%	2,293	46.84%	2,301	42.78%	2,290	42.19%	2,287	44.05%	2,287	44.05%	2,246	28.85%
Çankırı	2,197	50.18%	2,054	19.11%	1,993	23.87%	1,927	16.08%	2,149	46.84%	2,169	42.78%	2,160	42.19%	2,152	44.05%	2,152	44.05%	2,157	28.85%
Kırıkkale	2,460	50.18%	2,567	25.27%	2,514	7.16%	2,290	19.61%	2,394	46.84%	2,393	42.78%	2,381	42.19%	2,383	44.05%	2,383	44.05%	2,308	28.85%
Gümüşhane	2,612	60.40%	2,623	25.27%	2,530	7.16%	2,469	0.00%	2,498	46.84%	2,487	42.78%	2,474	42.19%	2,480	44.05%	2,480	44.05%	2,372	28.85%
Ardahan	2,497	0.00%	2,252	19.11%	2,551	7.16%	3,295	0.00%	2,635	46.84%	2,613	42.78%	2,598	42.19%	2,609	44.05%	2,609	44.05%	2,456	0.00%
Kilis	1,537	66.93%	1,909	19.11%	1,813	23.87%	1,805	16.08%	1,795	46.84%	1,316	42.78%	1,316	42.19%	1,316	44.05%	1,278	44.05%	1,278	0.00%
Bayburt	2,111	60.40%	2,413	25.27%	2,470	7.16%	2,171	19.61%	2,109	46.84%	2,133	42.78%	2,124	42.19%	2,114	44.05%	2,114	0.00%	2,132	0.00%
Hakkari	2,497	0.00%	2,643	25.27%	2,859	0.00%	2,859	0.00%	2,535	46.84%	2,522	42.78%	2,508	42.19%	2,859	0.00%	2,859	0.00%	2,859	0.00%
Tunceli	2,497	0.00%	2,825	25.27%	3,581	0.00%	3,581	0.00%	2,874	46.84%	2,831	42.78%	3,582	0.00%	3,582	0.00%	3,582	0.00%	3,582	0.00%

Table D.11: Bühlmann-Straub premiums and credibility factors - 1

	Region		K-means						Hierarchical											
City	$k = 7$		$k = 3$		$k = 6$		$k = 9$		$k = 2$		$k = 3$		$k = 4$		$k = 5$		$k = 6$		$k = 7$	
İstanbul	2,152	83.15%	2,162	0.00%	2,162	0.00%	2,162	0.00%	2,162	0.00%	2,162	0.00%	2,162	0.00%	2,162	0.00%	2,162	0.00%	2,162	0.00%
Ankara	2,037	0.00%	2,017	94.57%	2,003	82.67%	2,003	82.67%	2,023	98.01%	2,007	87.37%	2,007	87.37%	2,007	87.37%	2,003	82.67%	2,003	82.67%
İzmir	1,783	91.38%	1,790	91.62%	1,810	74.96%	1,810	74.96%	1,786	96.87%	1,801	81.28%	1,801	81.28%	1,801	81.28%	1,810	74.96%	1,810	74.96%
Bursa	2,018	37.27%	1,887	85.34%	1,882	82.37%	1,837	11.05%	1,890	94.28%	1,888	1.55%	1,888	1.55%	1,888	1.55%	1,888	95.81%	1,888	95.71%
Antalya	2,156	91.03%	2,129	81.72%	2,219	50.07%	2,120	0.00%	2,166	92.67%	2,169	95.94%	2,169	95.90%	2,169	95.81%	2,169	94.60%	2,168	94.48%
Adana	1,832	88.85%	1,840	77.83%	1,834	73.79%	1,830	6.97%	1,837	90.85%	1,832	94.77%	1,832	94.73%	1,832	94.60%	1,832	93.23%	1,831	93.07%
Kocaeli	2,150	20.51%	2,346	18.68%	2,294	36.19%	2,358	0.00%	2,310	87.73%	2,322	93.44%	2,321	93.38%	2,320	93.23%	2,320	90.84%	2,318	90.64%
Mersin	2,097	84.04%	2,063	69.86%	2,215	34.21%	2,120	0.00%	2,112	86.77%	2,117	91.11%	2,116	91.04%	2,116	90.84%	2,116	90.09%	2,114	89.88%
Gaziantep	1,711	86.74%	1,760	67.23%	1,758	62.21%	1,825	4.19%	1,734	85.30%	1,721	90.39%	1,721	90.31%	1,722	90.09%	1,722	88.95%	1,721	88.71%
Kayseri	2,037	0.00%	1,999	63.69%	1,977	58.45%	2,120	0.00%	2,043	83.22%	2,044	89.27%	2,044	89.19%	2,044	88.95%	2,044	87.31%	2,042	87.04%
Samsun	2,279	87.65%	2,338	12.85%	2,271	26.69%	2,358	0.00%	2,247	82.11%	2,261	87.67%	2,260	87.58%	2,259	87.31%	2,259	86.43%	2,256	86.14%
Balıkesir	2,101	13.88%	2,040	61.24%	2,013	55.90%	2,120	0.00%	2,102	81.71%	2,107	86.81%	2,107	86.71%	2,106	86.43%	2,106	86.11%	2,104	85.82%
Muğla	1,926	60.13%	1,964	60.86%	1,943	55.50%	1,837	3.21%	2,000	81.47%	1,999	86.50%	1,999	86.40%	1,999	86.11%	1,999	85.91%	1,997	85.62%
Denizli	1,830	60.05%	1,866	60.78%	1,854	55.42%	1,832	3.20%	1,869	81.42%	1,861	86.31%	1,861	86.21%	1,861	85.91%	1,861	85.87%	1,860	85.58%
Manisa	1,772	57.50%	1,810	58.24%	1,802	52.81%	1,829	2.89%	1,792	79.77%	1,778	86.27%	1,778	86.17%	1,778	85.87%	1,778	84.55%	1,777	84.23%
Aydın	1,756	56.90%	1,795	57.65%	1,788	52.20%	1,828	2.82%	1,771	79.38%	1,755	84.98%	1,755	84.87%	1,756	84.55%	1,756	84.23%	1,755	83.90%
Hatay	1,926	74.67%	1,919	56.48%	1,900	51.01%	1,834	2.69%	1,943	78.58%	1,938	84.66%	1,938	84.55%	1,938	84.23%	1,938	83.58%	1,936	83.25%
Eskişehir	2,037	0.00%	1,902	56.21%	1,885	50.74%	1,833	2.67%	1,920	78.40%	1,914	84.03%	1,914	83.92%	1,914	83.58%	1,914	83.43%	1,912	83.10%
Konya	2,037	0.00%	2,332	10.04%	2,253	21.60%	2,120	0.00%	2,177	77.65%	2,189	83.89%	2,188	83.77%	2,187	83.43%	2,187	82.81%	2,184	82.47%
Tekirdağ	2,080	10.98%	1,924	54.72%	1,905	49.23%	1,834	2.51%	1,952	77.37%	1,948	83.28%	1,948	83.16%	1,948	82.81%	1,948	82.58%	1,946	82.23%
Trabzon	2,124	83.82%	2,025	54.21%	2,232	20.99%	2,120	0.00%	2,095	77.00%	2,102	83.06%	2,101	82.93%	2,100	82.58%	2,100	82.28%	2,098	81.93%
Zonguldak	1,933	80.36%	1,898	48.34%	1,878	42.88%	1,832	1.96%	1,918	72.69%	1,910	82.76%	1,909	82.64%	1,909	82.28%	1,909	78.69%	1,907	78.28%
Sakarya	2,127	8.76%	2,351	7.88%	2,291	17.43%	2,358	0.00%	2,319	72.57%	2,347	79.24%	2,346	79.10%	2,344	78.69%	2,344	78.59%	2,339	78.18%
Erzurum	2,070	89.67%	1,984	46.53%	1,954	41.12%	2,120	0.00%	2,050	71.11%	2,054	79.14%	2,054	79.00%	2,053	78.59%	2,053	77.35%	2,050	76.92%
Diyarbakır	2,029	73.38%	1,991	46.35%	1,961	40.94%	2,120	0.00%	2,063	70.96%	2,068	77.92%	2,067	77.77%	2,066	77.35%	2,066	77.22%	2,063	76.79%
Ordu	2,235	74.94%	2,340	5.85%	2,262	13.30%	2,120	0.00%	2,179	65.91%	2,199	77.80%	2,198	77.65%	2,196	77.22%	2,196	72.84%	2,191	72.36%
Çanakkale	2,096	6.37%	1,972	40.00%	1,941	34.84%	2,120	0.00%	2,045	65.34%	2,049	73.49%	2,049	73.32%	2,048	72.84%	2,048	72.34%	2,044	71.85%
Kahramanmaraş	1,888	59.49%	1,891	39.26%	1,870	34.15%	1,832	1.36%	1,912	64.64%	1,901	73.00%	1,901	72.83%	1,901	72.34%	1,901	71.72%	1,898	71.23%
Şanlıurfa	2,056	65.50%	2,001	37.31%	2,248	11.78%	2,120	0.00%	2,098	62.88%	2,110	72.39%	2,109	72.21%	2,107	71.72%	2,107	70.15%	2,103	69.64%
Malatya	1,900	85.67%	1,894	37.46%	1,872	32.46%	1,832	1.26%	1,919	62.73%	1,908	70.84%	1,908	70.66%	1,908	70.15%	1,908	70.01%	1,904	69.50%
Edirne	2,085	5.71%	1,897	37.24%	1,874	32.26%	1,832	1.25%	1,924	62.66%	1,913	70.70%	1,913	70.53%	1,912	70.01%	1,912	69.95%	1,909	69.44%
Afyonkarahisar	1,859	35.64%	1,918	36.34%	1,893	31.41%	1,833	1.20%	1,961	61.75%	1,955	70.65%	1,955	70.47%	1,954	69.95%	1,954	69.13%	1,950	68.61%
Çorum	2,062	70.03%	1,952	34.82%	1,921	30.00%	2,120	0.00%	2,021	60.17%	2,023	69.83%	2,022	69.65%	2,021	69.13%	2,021	67.70%	2,017	67.17%
Elazığ	2,035	83.15%	1,948	33.09%	1,917	28.41%	2,120	0.00%	2,017	58.31%	2,018	68.42%	2,017	68.23%	2,016	67.70%	2,016	65.99%	2,012	65.44%
Giresun	1,983	67.88%	1,910	32.58%	1,885	27.94%	1,832	1.02%	1,952	57.74%	1,943	66.73%	1,943	66.54%	1,942	65.99%	1,942	65.47%	1,938	64.92%
Sivas	2,037	0.00%	1,948	32.20%	1,916	27.59%	2,120	0.00%	2,018	57.32%	2,020	66.21%	2,019	66.02%	2,018	65.47%	2,018	65.07%	2,013	64.52%
Batman	1,797	59.25%	1,858	31.31%	1,839	26.78%	1,830	0.96%	1,859	56.32%	1,836	65.82%	1,836	65.63%	1,837	65.07%	1,837	64.14%	1,833	63.58%
Kırklareli	2,083	4.18%	1,862	29.94%	1,842	25.53%	1,831	0.90%	1,865	54.72%	1,842	64.90%	1,842	64.70%	1,843	64.14%	1,843	62.63%	1,839	62.06%
Tokat	2,115	65.10%	1,968	29.89%	1,933	25.49%	2,120	0.00%	2,061	54.66%	2,069	63.41%	2,068	63.21%	2,067	62.63%	2,067	62.58%	2,061	62.02%
Uşak	1,798	28.05%	1,863	28.67%	1,843	24.39%	1,831	0.85%	1,868	53.20%	1,845	63.36%	1,845	63.16%	1,845	62.58%	1,845	61.19%	1,842	60.62%

Table D.12: Bühlmann premiums and credibility factors - 2

	Region		K-means						Hierarchical											
City	$k=7$		$k=3$		$k=6$		$k=9$		$k=2$		$k=3$		$k=4$		$k=5$		$k=6$		$k=7$	
Mardin	2,040	55.70%	2,338	3.46%	2,254	8.12%	2,120	0.00%	2,094	52.71%	2,108	61.98%	2,107	61.78%	2,105	61.19%	2,105	60.73%	2,099	60.15%
Kütahya	1,930	26.25%	2,341	3.23%	2,260	7.60%	2,358	0.00%	2,128	51.12%	2,149	61.52%	2,148	61.31%	2,146	60.73%	2,146	59.20%	2,139	58.61%
Van	2,460	78.68%	2,351	3.25%	2,284	7.66%	2,358	0.00%	2,290	50.92%	2,339	59.99%	2,337	59.79%	2,333	59.20%	2,333	59.00%	2,324	58.42%
Isparta	1,888	45.11%	1,891	26.57%	1,866	22.50%	1,831	0.77%	1,921	50.58%	1,906	59.80%	1,905	59.60%	1,905	59.00%	1,905	58.67%	1,901	58.08%
Düzce	2,306	61.08%	2,346	3.16%	2,272	7.45%	2,358	0.00%	2,213	50.36%	2,249	59.47%	2,247	59.26%	2,244	58.67%	2,244	58.46%	2,236	57.87%
Osmaniye	1,800	44.86%	1,839	26.37%	1,822	22.32%	1,830	0.76%	1,823	50.32%	1,790	59.26%	1,790	59.06%	1,791	58.46%	1,791	58.42%	1,788	57.84%
Adıyaman	1,838	52.47%	1,883	25.71%	1,860	21.73%	1,831	0.73%	1,908	49.46%	1,889	59.23%	1,889	59.02%	1,889	58.42%	1,889	57.59%	1,885	56.99%
Rize	2,387	59.62%	2,351	2.98%	2,282	7.04%	2,524	34.86%	2,276	48.84%	2,325	58.40%	2,324	58.19%	2,319	57.59%	2,319	56.98%	2,310	56.39%
Kastamonu	2,020	58.05%	1,917	24.04%	1,887	20.25%	1,832	0.67%	1,975	47.35%	1,968	57.79%	1,968	57.58%	1,967	56.98%	1,967	55.51%	1,961	54.91%
Amasya	2,174	58.17%	1,981	24.13%	2,257	6.66%	2,120	0.00%	2,100	47.22%	2,117	56.32%	2,116	56.11%	2,113	55.51%	2,113	55.38%	2,106	54.78%
Yalova	2,095	2.96%	1,934	23.04%	1,902	19.36%	2,120	0.00%	2,011	45.84%	2,011	56.20%	2,011	55.99%	2,009	55.38%	2,009	54.01%	2,003	53.41%
Aksaray	2,037	0.00%	2,343	2.19%	2,262	5.25%	2,358	0.00%	2,124	41.38%	2,149	54.83%	2,148	54.62%	2,145	54.01%	2,145	49.47%	2,137	48.87%
Nevşehir	2,037	0.00%	2,342	2.22%	2,260	5.30%	2,120	0.00%	2,104	41.12%	2,124	50.30%	2,123	50.08%	2,120	49.47%	2,120	49.20%	2,112	48.60%
Niğde	2,037	0.00%	1,963	19.63%	2,256	5.19%	2,306	2.06%	2,077	40.86%	2,092	50.03%	2,090	49.82%	2,088	49.20%	2,088	48.93%	2,080	48.33%
Bolu	2,310	51.08%	2,347	2.12%	2,272	5.08%	2,120	0.00%	2,197	40.30%	2,238	49.76%	2,237	49.55%	2,232	48.93%	2,232	48.36%	2,223	47.75%
Karabük	2,028	48.28%	1,911	17.59%	1,880	14.62%	1,832	0.45%	1,972	37.64%	1,964	49.19%	1,963	48.97%	1,962	48.36%	1,962	45.57%	1,956	44.97%
Siirt	1,966	37.64%	1,944	15.91%	1,907	13.18%	1,832	0.40%	2,047	34.85%	2,056	46.39%	2,055	46.18%	2,053	45.57%	2,053	42.60%	2,045	42.01%
Yozgat	2,037	0.00%	2,354	1.62%	2,287	3.90%	2,319	1.53%	2,320	33.98%	2,398	43.41%	2,395	43.20%	2,388	42.60%	2,388	41.66%	2,896	41.07%
Bilecik	2,104	1.82%	2,346	1.63%	2,267	3.92%	2,358	0.00%	2,151	33.86%	2,187	42.47%	2,185	42.26%	2,181	41.66%	2,181	41.53%	2,171	39.75%
Şırnak	1,993	35.47%	2,342	1.54%	2,259	3.72%	2,120	0.00%	2,074	32.76%	2,091	42.34%	2,089	42.13%	2,087	41.53%	2,087	40.33%	2,078	39.75%
Burdur	2,105	28.13%	2,347	1.54%	2,272	3.72%	2,358	0.00%	2,186	32.76%	2,231	41.13%	2,229	40.92%	2,225	40.33%	2,225	40.33%	2,214	38.61%
Ağrı	2,108	62.11%	1,939	14.11%	2,256	3.55%	2,120	0.00%	2,042	31.72%	2,051	41.13%	2,050	40.92%	2,047	40.33%	2,047	39.19%	2,039	37.43%
Artvin	2,377	40.59%	2,350	1.40%	2,277	3.39%	2,358	0.00%	2,230	30.64%	2,288	39.98%	2,286	39.77%	2,280	39.19%	2,280	38.00%	2,268	36.55%
Sinop	2,295	39.69%	2,347	1.35%	2,270	3.26%	2,358	0.00%	2,166	29.85%	2,208	38.78%	2,206	38.58%	2,201	38.00%	2,201	37.12%	2,190	36.32%
Erzincan	2,330	59.78%	2,346	1.34%	2,268	3.23%	2,502	19.10%	2,150	29.64%	2,188	37.90%	2,187	37.69%	2,182	37.12%	2,182	36.88%	2,171	35.74%
Kırşehir	2,037	0.00%	1,931	12.68%	1,895	10.44%	2,120	0.00%	2,027	29.12%	2,032	37.66%	2,031	37.46%	2,029	36.88%	2,029	36.30%	2,020	33.20%
Kars	2,101	56.43%	1,933	11.49%	1,896	9.43%	2,307	1.10%	2,034	26.85%	2,041	37.07%	2,040	36.87%	2,038	36.30%	2,038	33.74%	2,029	31.67%
Bitlis	2,344	54.70%	2,346	1.09%	2,268	2.64%	2,358	0.00%	2,146	25.50%	2,186	34.48%	2,184	34.29%	2,180	33.74%	2,180	32.19%	2,168	30.93%
Karaman	2,037	0.00%	1,916	10.42%	1,882	8.54%	1,832	0.25%	1,997	24.85%	1,994	32.92%	1,993	32.73%	1,991	32.19%	1,991	31.45%	1,983	30.82%
Muş	2,264	53.85%	2,345	1.05%	2,265	2.56%	2,491	15.64%	2,108	24.76%	2,137	32.17%	2,135	31.98%	2,131	31.45%	2,131	31.34%	2,120	29.41%
Bartın	2,450	32.99%	2,352	1.01%	2,281	2.46%	2,594	15.14%	2,269	24.14%	2,346	32.06%	2,343	31.87%	2,336	31.34%	2,336	30.63%	2,896	28.39%
Bingöl	2,311	52.06%	2,346	0.98%	2,267	2.38%	2,358	0.00%	2,126	23.53%	2,161	31.34%	2,159	31.15%	2,155	30.63%	2,155	29.92%	2,143	26.74%
Iğdır	2,246	50.81%	2,345	0.93%	2,264	2.27%	2,120	0.00%	2,095	22.65%	2,121	30.62%	2,120	30.44%	2,116	29.92%	2,116	28.88%	2,105	24.43%
Çankırı	2,037	0.00%	1,920	8.70%	1,885	7.10%	1,832	0.20%	2,010	21.23%	2,011	29.57%	2,010	29.39%	2,007	28.88%	2,007	27.21%	1,998	19.86%
Kırıkkale	2,037	0.00%	2,345	0.76%	2,265	1.86%	2,120	0.00%	2,100	19.27%	2,130	27.87%	2,128	27.70%	2,124	27.21%	2,124	24.88%	2,112	17.90%
Gümüşhane	2,275	22.06%	2,347	0.58%	2,268	1.43%	2,358	0.00%	2,125	15.47%	2,165	25.51%	2,163	25.34%	2,158	24.88%	2,158	20.25%	2,144	12.29%
Ardahan	2,190	36.23%	1,935	5.39%	2,262	1.26%	2,309	0.49%	2,057	13.87%	2,073	20.79%	2,072	20.65%	2,068	20.25%	2,068	18.26%	2,056	0.00%
Kilis	1,836	11.22%	1,896	3.81%	1,864	3.08%	1,831	0.08%	1,960	10.07%	1,887	18.76%	1,887	18.63%	1,887	18.26%	1,181	12.56%	1,181	0.00%
Bayburt	2,180	13.80%	2,345	0.33%	2,262	0.81%	2,120	0.00%	2,050	9.38%	2,065	12.92%	2,064	12.83%	2,060	12.56%	2,060	0.00%	2,048	0.00%
Hakkari	2,529	26.56%	2,349	0.33%	3,348	0.00%	3,348	0.00%	2,168	9.30%	2,228	12.82%	2,225	12.72%	3,348	0.00%	3,348	0.00%	3,348	0.00%
Tunceli	2,374	17.23%	2,347	0.19%	3,051	0.00%	3,051	0.00%	2,103	5.57%	2,141	7.80%	3,051	0.00%	3,051	0.00%	3,051	0.00%	3,051	0.00%

All comparisons are made in light of the 2011 Q1 results given below.

Table D.13: 2011 Q1 claim averages - (target premium)

2011-Q1 (Accident Year-Quarter) Claim Averages by Cities					
İstanbul	2,016	Kahramanmaraş	1,641	Bolu	1,744
Ankara	1,906	Şanlıurfa	1,746	Karabük	1,467
İzmir	1,628	Malatya	1,643	Siirt	1,548
Bursa	1,781	Edirne	1,618	Yozgat	1,897
Antalya	1,867	Afyonkarahisar	2,269	Bilecik	1,616
Adana	1,745	Çorum	2,258	Şırnak	1,481
Kocaeli	2,284	Elazığ	1,548	Burdur	2,580
Mersin	1,995	Giresun	1,476	Ağrı	3,200
Gaziantep	1,617	Sivas	1,923	Artvin	2,581
Kayseri	2,032	Batman	1,372	Sinop	2,025
Samsun	2,029	Kırklareli	1,581	Erzincan	1,606
Balıkesir	1,927	Tokat	1,944	Kırşehir	1,725
Muğla	1,989	Uşak	1,768	Kars	1,839
Denizli	1,562	Mardin	2,393	Bitlis	2,161
Manisa	1,609	Kütahya	1,933	Karaman	2,165
Aydın	1,516	Van	2,155	Muş	1,321
Hatay	1,620	Isparta	1,576	Bartın	2,187
Eskişehir	1,839	Düzce	2,327	Bingöl	2,276
Konya	1,985	Osmaniye	1,389	Iğdır	5,173
Tekirdağ	1,695	Adıyaman	1,737	Çankırı	1,852
Trabzon	2,036	Rize	1,948	Kırıkkale	1,613
Zonguldak	1,581	Kastamonu	1,606	Gümüşhane	1,852
Sakarya	2,098	Amasya	1,713	Ardahan	2,280
Erzurum	1,880	Yalova	1,991	Kilis	705
Diyarbakır	1,808	Aksaray	2,188	Bayburt	2,412
Ordu	1,906	Nevşehir	1,703	Hakkari	2,552
Çanakkale	1,665	Niğde	2,523	Tunceli	2,108