Project

Welcome to my project page.

Keep an eye on this space to stay updated with my project activities.

# 1. Project Overview and Scope 🔎

Istanbul is not only Turkey’s but also Europe’s most populous city. Therefore, if a disaster such as an earthquake occurs, the potential of destruction and the need for help would be very high. Due to the earthquake disaster that occurred at the beginning of 2023, the risky condition of buildings and residences in Turkey has become a matter of concern for people.

The purpose of this project is to examine Istanbul, which is almost certain to face a possible earthquake and significant destruction. The aim of the project is not only to assess buildings and residences in terms of their risky structural condition and make district-based interpretations but also to evaluate the opinions of people living in Istanbul and compare all the results.

7 datasets were used for this purpose, various graphs were drawn and interpreted comprehensively to make inferences.

# 2. Data 📄

## 2.1 Data Source

* [Deprem Senaryosu Analiz Sonuçları](https://data.ibb.gov.tr/dataset/deprem-senaryosu-analiz-sonuclari/resource/9c3ac492-de4b-4245-b418-7ad3df67a193)
* [2017 Yılı Mahalle Bazlı Bina Sayıları](https://data.ibb.gov.tr/dataset/mahalle-bazli-bina-analiz-verisi/resource/cef193d5-0bd2-4e8d-8a69-275c50288875)
* [VDYM İlçe Bazında Hanelerdeki Riskli Yapı Durumu Bireysel Öngörüleri](https://data.ibb.gov.tr/dataset/vdym-ilce-bazinda-hanelerdeki-riskli-yapi-durumu-bireysel-ongoruleri/resource/52fe7088-62f2-4480-8f27-788dee385c9b)
* [VDYM İlçe Bazında Hanelerdeki Kentsel Dönüşüm Fikri Bireysel Öngörüleri](https://data.ibb.gov.tr/dataset/vdym-ilce-bazinda-hanelerdeki-kentsel-donusum-fikri-bireysel-ongoruleri/resource/2683ee8b-9bc3-45cf-a115-49b285851317)
* [İlçe Bazlı Ortalama Hane Halkı Büyüklüğü](https://data.ibb.gov.tr/dataset/ilce-bazli-ortalama-hane-halki-buyuklugu/resource/da747beb-df41-474e-aa8a-93d983e69b15)
* [2019 Yılı Belediye Nüfusları](https://data.ibb.gov.tr/dataset/belediye-nufuslari-veri-seti/resource/c6c9b289-2824-41b3-ab3d-4fd655ed4e24)
* [İstanbul Çevresinde Gerçekleşen Depremler](https://data.ibb.gov.tr/dataset/istanbul-da-son-bir-yilda-gerceklesen-depremler/resource/32904f60-8091-4dc3-b527-351dca6c1c22)

Information and explanations about these datasets are included in the section [2.2 General Information About Data](#general-information-about-data).

All this data was taken from [İstanbul Büyükşehir Belediyesi Açık Veri Portalı](https://data.ibb.gov.tr/dataset) and [TÜİK Merkezi Dağıtım Sistemi](https://biruni.tuik.gov.tr/medas/).

#install.packages("readxl")  
library(readxl)  
  
dosya\_yolu<-"/Users/aycacetin/Documents/GitHub/emu660-spring2024-aycacetin/assets/deprem.xlsx"  
  
deprem\_senaryo<-read\_excel(dosya\_yolu, sheet=1)  
bina\_sayi<-read\_excel(dosya\_yolu, sheet=2)  
riskli\_yapi\_durumu<-read\_excel(dosya\_yolu, sheet=3)  
kentsel\_donusum\_fikri<-read\_excel(dosya\_yolu, sheet=4)  
ort\_hane<-read\_excel(dosya\_yolu, sheet=7)  
belediye\_nufus<-read\_excel(dosya\_yolu, sheet=8)  
gerceklesen\_deprem<-read\_excel(dosya\_yolu, sheet=9)  
  
save(deprem\_senaryo, file="deprem\_senaryo.RData")  
save(bina\_sayi, file="bina\_sayi.RData")  
save(riskli\_yapi\_durumu, file="riskli\_yapi\_durumu.RData")  
save(kentsel\_donusum\_fikri, file="kentsel\_donusum\_fikri.RData")  
save(ort\_hane, file="ort\_hane.RData")  
save(belediye\_nufus, file="belediye\_nufus.RData")  
save(gerceklesen\_deprem, file="gerceklesen\_deprem.RData")  
  
str(deprem\_senaryo)

tibble [959 × 16] (S3: tbl\_df/tbl/data.frame)  
 $ id : num [1:959] 1 2 3 4 5 6 7 8 9 10 ...  
 $ ilce\_adi : chr [1:959] "ADALAR" "ADALAR" "ADALAR" "ADALAR" ...  
 $ mahalle\_adi : chr [1:959] "BURGAZADA" "HEYBELİADA" "KINALIADA" "MADEN" ...  
 $ mahalle\_koy\_uavt : num [1:959] 40139 40142 40143 40140 40141 ...  
 $ cok\_agir\_hasarli\_bina\_sayisi: num [1:959] 54 101 53 104 101 1 2 1 1 1 ...  
 $ agir\_hasarli\_bina\_sayisi : num [1:959] 99 175 97 192 180 3 13 6 6 5 ...  
 $ orta\_hasarli\_bina\_sayisi : num [1:959] 256 423 287 483 445 21 108 51 49 23 ...  
 $ hafif\_hasarli\_bina\_sayisi : num [1:959] 241 393 302 484 422 57 371 199 129 65 ...  
 $ can\_kaybi\_sayisi : num [1:959] 8 25 5 22 16 0 0 0 0 0 ...  
 $ agir\_yarali\_sayisi : num [1:959] 6 21 3 18 13 0 0 0 0 0 ...  
 $ hastanede\_tedavi\_sayisi : num [1:959] 24 66 15 64 48 1 8 2 2 0 ...  
 $ hafif\_yarali\_sayisi : num [1:959] 42 113 27 113 83 3 26 10 8 0 ...  
 $ dogalgaz\_boru\_hasari : num [1:959] 0 1 0 1 1 0 1 0 0 0 ...  
 $ icme\_suyu\_boru\_hasari : num [1:959] 0 1 1 1 2 0 1 1 0 0 ...  
 $ atik\_su\_boru\_hasari : num [1:959] 1 2 1 2 2 0 1 1 0 0 ...  
 $ gecici\_barinma : num [1:959] 398 763 420 847 687 89 659 273 209 26 ...

str(bina\_sayi)

tibble [959 × 10] (S3: tbl\_df/tbl/data.frame)  
 $ id : num [1:959] 1 2 3 4 5 6 7 8 9 10 ...  
 $ ilce\_adi : chr [1:959] "ADALAR" "ADALAR" "ADALAR" "ADALAR" ...  
 $ mahalle\_adi : chr [1:959] "BURGAZADA" "HEYBELİADA" "KINALIADA" "MADEN" ...  
 $ mahalle\_uavt : num [1:959] 40139 40142 40143 40140 40141 ...  
 $ 1980\_oncesi : num [1:959] 433 836 610 863 842 0 0 0 0 0 ...  
 $ 1980-2000\_arasi: num [1:959] 214 347 244 510 426 244 1360 685 565 332 ...  
 $ 2000\_sonrasi : num [1:959] 173 212 158 308 217 121 845 589 216 184 ...  
 $ 1-4 kat\_arasi : num [1:959] 802 1359 923 1637 1434 ...  
 $ 5-9 kat\_arasi : num [1:959] 18 36 89 44 51 12 555 470 27 1 ...  
 $ 9-19 kat\_arasi : num [1:959] 0 0 0 0 0 0 3 7 0 0 ...

str(riskli\_yapi\_durumu)

tibble [39 × 4] (S3: tbl\_df/tbl/data.frame)  
 $ ilce : chr [1:39] "ADALAR" "ARNAVUTKÖY" "ATAŞEHİR" "AVCILAR" ...  
 $ riskli\_yapi\_evet : num [1:39] 28 96 384 111 508 470 227 103 170 118 ...  
 $ riskli\_yapi\_hayir : num [1:39] 295 603 910 971 1176 ...  
 $ riskli\_yapi\_bilmiyorum: num [1:39] 36 84 102 307 350 284 114 57 129 140 ...

str(kentsel\_donusum\_fikri)

tibble [39 × 3] (S3: tbl\_df/tbl/data.frame)  
 $ ilce : chr [1:39] "ADALAR" "ARNAVUTKÖY" "ATAŞEHİR" "AVCILAR" ...  
 $ kentsel\_donusum\_ihtiyac\_evet : num [1:39] 40 100 457 240 607 536 263 59 190 158 ...  
 $ kentsel\_donusum\_ihtiyac\_hayir: num [1:39] 319 683 939 1148 1427 ...

str(ort\_hane)

tibble [39 × 2] (S3: tbl\_df/tbl/data.frame)  
 $ ilce : chr [1:39] "Adalar" "Arnavutköy" "Ataşehir" "Avcılar" ...  
 $ ortalama\_hanehalki\_buyuklugu: num [1:39] 2.37 3.89 2.99 3.29 3.77 3.32 2.8 3.65 3.28 2.36 ...

str(belediye\_nufus)

tibble [39 × 2] (S3: tbl\_df/tbl/data.frame)  
 $ belediye : chr [1:39] "Adalar Belediyesi" "Arnavutköy Belediyesi" "Ataşehir Belediyesi" "Avcılar Belediyesi" ...  
 $ nufus\_2019: num [1:39] 15238 282488 425094 448882 745125 ...

str(gerceklesen\_deprem)

tibble [349 × 8] (S3: tbl\_df/tbl/data.frame)  
 $ event\_id : chr [1:349] "20200204\_0000090" "20200204\_0000082" "20200204\_0000067" "20200204\_0000046" ...  
 $ time : chr [1:349] "2020-02-04T10:17:09.4Z" "2020-02-04T08:43:16.0Z" "2020-02-04T06:13:29.2Z" "2020-02-04T03:42:21.5Z" ...  
 $ latitude : num [1:349] 39 39 38.4 38.4 39.1 ...  
 $ longitude : num [1:349] 27.9 27.9 25.5 25.5 27.8 ...  
 $ depth/km : num [1:349] 6 10 15 17 9 12 10 2 7 9 ...  
 $ mag\_type : chr [1:349] "ml" "ml" "ml" "ml" ...  
 $ magnitude : num [1:349] 3.4 4.1 3.3 3.3 3.3 3.4 3.4 3.3 4 3.5 ...  
 $ event\_location\_name: chr [1:349] "WESTERN TURKEY" "WESTERN TURKEY" "AEGEAN SEA" "AEGEAN SEA" ...

## 2.2 General Information About Data

[Deprem Senaryosu Analiz Sonuçları](https://data.ibb.gov.tr/dataset/deprem-senaryosu-analiz-sonuclari/resource/9c3ac492-de4b-4245-b418-7ad3df67a193) is a dataset containing the results of an analysis made on an earthquake scenario that is predicted to have a magnitude of 7.5. Each row represents a neighbourhood. The columns have information mainly about the number of buildings based on damage level, the number of people based on vital status and pipe damages.

[2017 Yılı Mahalle Bazlı Bina Sayıları](https://data.ibb.gov.tr/dataset/mahalle-bazli-bina-analiz-verisi/resource/cef193d5-0bd2-4e8d-8a69-275c50288875) is a dataset about the buildings. Each row represents a neighbourhood. The columns have information mainly about year of construction and number of floors of the buildings.

[VDYM İlçe Bazında Hanelerdeki Riskli Yapı Durumu Bireysel Öngörüleri](https://data.ibb.gov.tr/dataset/vdym-ilce-bazinda-hanelerdeki-riskli-yapi-durumu-bireysel-ongoruleri/resource/52fe7088-62f2-4480-8f27-788dee385c9b) is a dataset about the opinions of households about the risky building situation on a district basis.

[VDYM İlçe Bazında Hanelerdeki Kentsel Dönüşüm Fikri Bireysel Öngörüleri](https://data.ibb.gov.tr/dataset/vdym-ilce-bazinda-hanelerdeki-kentsel-donusum-fikri-bireysel-ongoruleri/resource/2683ee8b-9bc3-45cf-a115-49b285851317) is a dataset about the opinions of households about the idea of urban transformation on a district basis.

[İlçe Bazlı Ortalama Hane Halkı Büyüklüğü](https://data.ibb.gov.tr/dataset/ilce-bazli-ortalama-hane-halki-buyuklugu/resource/da747beb-df41-474e-aa8a-93d983e69b15) is a dataset about the average household size on a district basis.

[2019 Yılı Belediye Nüfusları](https://data.ibb.gov.tr/dataset/belediye-nufuslari-veri-seti/resource/c6c9b289-2824-41b3-ab3d-4fd655ed4e24) is a dataset about the population within the borders of each municipality.

[İstanbul Çevresinde Gerçekleşen Depremler](https://data.ibb.gov.tr/dataset/istanbul-da-son-bir-yilda-gerceklesen-depremler/resource/32904f60-8091-4dc3-b527-351dca6c1c22) is a dataset about the earthquakes around Istanbul within 1 year. Each row represents an earthquake. In the columns, there are details about the earthquakes.

## 2.3 Reason of Choice

Istanbul is a city where the risk of earthquake is said to be very high, and it is predicted to cause significant loss of life and property. Almost every day, warnings from experts about the earthquake, which is estimated to be between 7.2 and 7.6 in magnitude, can be seen in news and newspapers. I decided to work on such a current topic, to inform people through comprehensive analysis, to make them more aware, and to encourage taking steps against this disaster to get through with minimal damage.

## 2.4 Preprocessing

There are 7 datasets in “.csv”, “.xls” and “.xlsx” formats which were downloaded from the sources mentioned in the section [2.1 Data Source](#data-source). All these datasets have been merged into the same Excel file called “deprem.xlsx” and organized for a better use in the analysis. Each dataset is placed on a different sheet. After that, spelling errors on each dataset were corrected via Excel. In other words, correction of wrong letters was done. Finally, the “deprem.xlsx” file was read and stored in [.RData](https://github.com/emu-hacettepe-analytics/emu660-spring2024-aycacetin/tree/main) format using the code in the section [2.1 Data Source](#data-source). The [.RData](https://github.com/emu-hacettepe-analytics/emu660-spring2024-aycacetin/tree/main) file can be downloaded to review.

# 3. Analysis 🖥️

## 3.1 Exploratory Data Analysis

### Analysis of the Datasets

#install.packages("ggplot2")  
library(ggplot2)  
library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':  
  
 filter, lag

The following objects are masked from 'package:base':  
  
 intersect, setdiff, setequal, union

library(tidyverse)

── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
✔ forcats 1.0.0 ✔ stringr 1.5.1  
✔ lubridate 1.9.3 ✔ tibble 3.2.1  
✔ purrr 1.0.2 ✔ tidyr 1.3.1  
✔ readr 2.1.5

── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
✖ dplyr::filter() masks stats::filter()  
✖ dplyr::lag() masks stats::lag()  
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

**“deprem\_senaryo” analysis:**

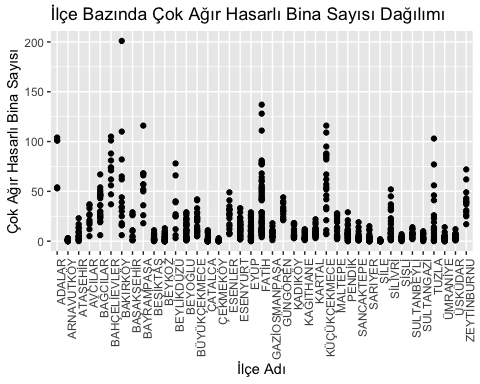
The analysis began with the “deprem\_senaryo” dataset. Initially, plots were created using data about the number of damaged buildings, the number of injured people and the number of pipeline damages. And then, the need for temporary shelter was assesed.

Although the data is based on neighborhoods, the analyses have been done based on districts.

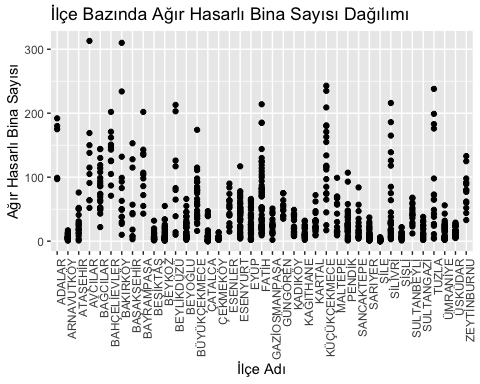
“ilçe bazında hasarlı bina sayıları” (4 types of damage):

* Bakırköy, Fatih and Küçükçekmece on the “Çok Ağır Hasarlı Binalar” plot,
* Avıclar and Bakırköy on the “Ağır Hasarlı Binalar” plot,
* Avcılar on the “Orta Hasarlı Binalar” and “Hafif Hasarlı Binalar” plot were attention-grabbing.

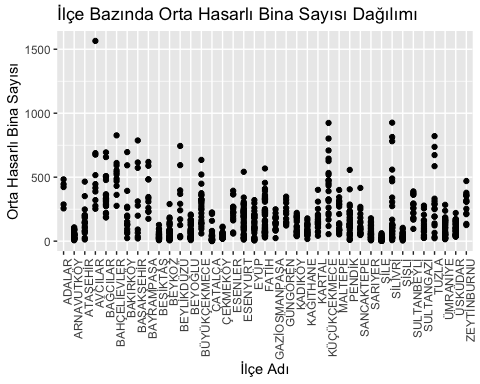
p1<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=cok\_agir\_hasarli\_bina\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Çok Ağır Hasarlı Bina Sayısı")+  
 labs(title="İlçe Bazında Çok Ağır Hasarlı Bina Sayısı Dağılımı")  
p1



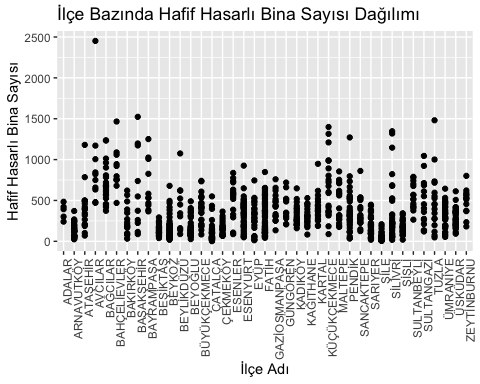
p2<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=agir\_hasarli\_bina\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Ağır Hasarlı Bina Sayısı")+  
 labs(title="İlçe Bazında Ağır Hasarlı Bina Sayısı Dağılımı")  
p2



p3<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=orta\_hasarli\_bina\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Orta Hasarlı Bina Sayısı")+  
 labs(title="İlçe Bazında Orta Hasarlı Bina Sayısı Dağılımı")  
p3



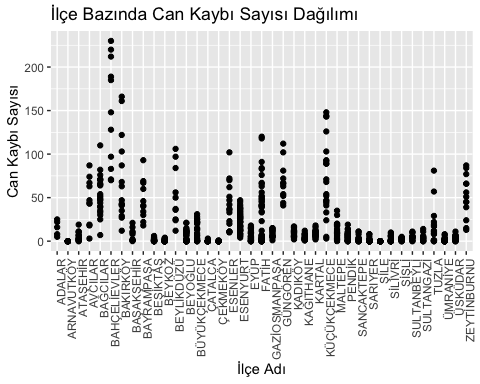
p4<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=hafif\_hasarli\_bina\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Hafif Hasarlı Bina Sayısı")+  
 labs(title="İlçe Bazında Hafif Hasarlı Bina Sayısı Dağılımı")  
p4



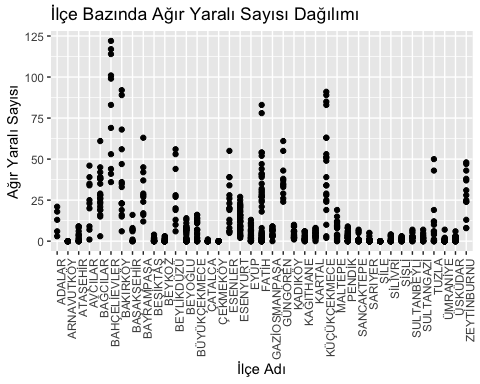
“kişi sayıları” (4 types):

* Bahçelievler and Bakırköy on the “Can Kaybı” plot,
* Bahçelievler, Bakırköy, Fatih and Küçükçekmece on the “Ağır Yaralı” plot,
* Bahçelievler, Bakırköy and Küçükçekmece on the “Hastanede Tedavi” and “Hafif Yaralı” plot were attention-grabbing.

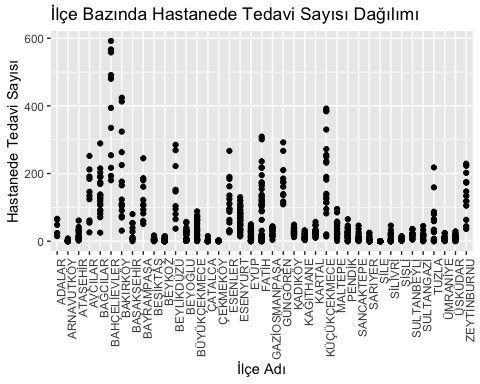
p5<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=can\_kaybi\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Can Kaybı Sayısı")+  
 labs(title="İlçe Bazında Can Kaybı Sayısı Dağılımı")  
p5



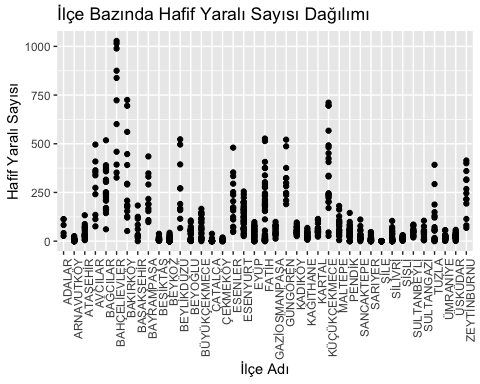
p6<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=agir\_yarali\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Ağır Yaralı Sayısı")+  
 labs(title="İlçe Bazında Ağır Yaralı Sayısı Dağılımı")  
p6



p7<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=hastanede\_tedavi\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Hastanede Tedavi Sayısı")+  
 labs(title="İlçe Bazında Hastanede Tedavi Sayısı Dağılımı")  
p7



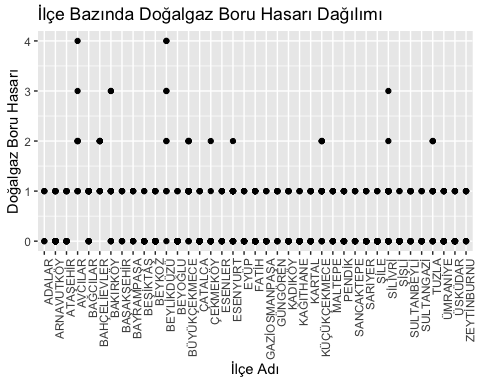
p8<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=hafif\_yarali\_sayisi))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Hafif Yaralı Sayısı")+  
 labs(title="İlçe Bazında Hafif Yaralı Sayısı Dağılımı")  
p8



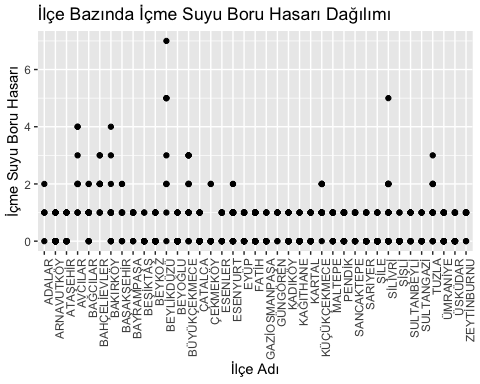
“boru hasarı” (3 types of damage):

* Avcılar and Beylikdüzü on the “Doğalgaz Boru Hasarı” plot,
* Beylikdüzü on the “İçme Suyu Boru Hasarı” plot,
* Avcılar, Bakırköy, Beylikdüzü and Tuzla on the “Atık Su Boru Hasarı” plot were attention-grabbing.

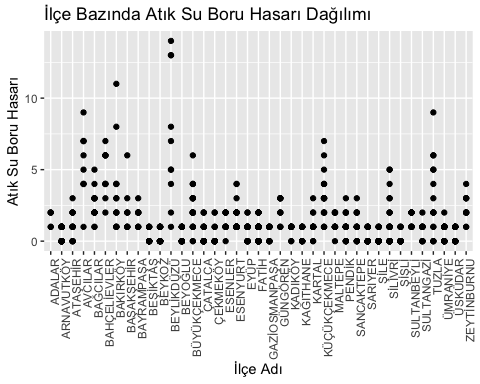
p9<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=dogalgaz\_boru\_hasari))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Doğalgaz Boru Hasarı")+  
 labs(title="İlçe Bazında Doğalgaz Boru Hasarı Dağılımı")  
p9



p10<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=icme\_suyu\_boru\_hasari))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("İçme Suyu Boru Hasarı")+  
 labs(title="İlçe Bazında İçme Suyu Boru Hasarı Dağılımı")  
p10



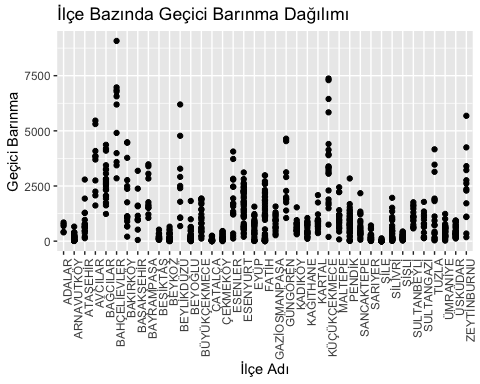
p11<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=atik\_su\_boru\_hasari))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Atık Su Boru Hasarı")+  
 labs(title="İlçe Bazında Atık Su Boru Hasarı Dağılımı")  
p11



“geçici barınma”:

* Avcılar, Bahçelievler, Beylikdüzü, Küçükçekmece and Zeytinburnu on the “Atık Su Boru Hasarı” plot were attention-grabbing.

p12<-deprem\_senaryo%>%ggplot(aes(x=ilce\_adi, y=gecici\_barinma))+  
 geom\_point()+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Geçici Barınma")+  
 labs(title="İlçe Bazında Geçici Barınma Dağılımı")  
p12



**“bina\_sayi” analysis:**

The analysis continued with the “bina\_sayi” dataset, evaluating buildings based on their construction year and number of floors.

Although the data is based on neighborhoods, the analyses have been done based on districts.

“ilçe bazında yıllara göre bina sayılarının oranları”:

It has been observed that the majority of the buildings were constructed between 1980 and 2000.

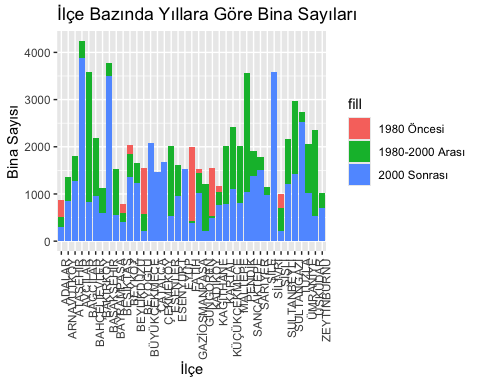
once<-sum(bina\_sayi$`1980\_oncesi`)  
ara<-sum(bina\_sayi$`1980-2000\_arasi`)  
sonra<-sum(bina\_sayi$`2000\_sonrasi`)  
oranlar1<-c(once=once/(once+ara+sonra),   
 ara=ara/(once+ara+sonra),   
 sonra=sonra/(once+ara+sonra))  
print(oranlar1)

once ara sonra   
0.2215122 0.4700309 0.3084569

“ilçe bazında yıllara göre bina sayıları”:

* Beyoğlu, Fatih and Kadıköy on the “Yıllara Göre Bina Sayıları” plot were attention-grabbing.

p13<-ggplot(bina\_sayi, aes(x=ilce\_adi))+  
 geom\_col(aes(y=`1980\_oncesi`, fill="1980 Öncesi"), position=position\_dodge(width=0.9))+  
 geom\_col(aes(y=`1980-2000\_arasi`, fill="1980-2000 Arası"), position=position\_dodge(width=0.9))+  
 geom\_col(aes(y=`2000\_sonrasi`, fill="2000 Sonrası"), position=position\_dodge(width=0.9))+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe")+  
 ylab("Bina Sayısı")+  
 labs(title="İlçe Bazında Yıllara Göre Bina Sayıları")  
p13



“ilçe bazında kat sayısına göre bina sayılarının oranları”:

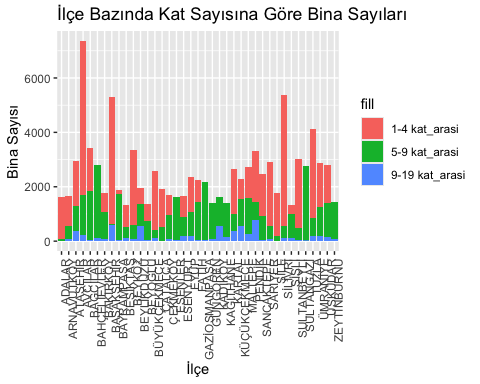
It has been observed that the majority of the buildings are 1-4 stories tall.

alcak<-sum(bina\_sayi$`1-4 kat\_arasi`)  
orta<-sum(bina\_sayi$`5-9 kat\_arasi`)  
yuksek<-sum(bina\_sayi$`9-19 kat\_arasi`)  
oranlar2<-c(alcak=alcak/(alcak+orta+yuksek),   
 orta=orta/(alcak+orta+yuksek),   
 yuksek=yuksek/(alcak+orta+yuksek))  
print(oranlar2)

alcak orta yuksek   
0.65751122 0.31906640 0.02342238

“ilçe bazında kat sayısına göre bina sayıları”:

p14<-ggplot(bina\_sayi, aes(x=ilce\_adi))+  
 geom\_col(aes(y=`1-4 kat\_arasi`, fill="1-4 kat\_arasi"), position=position\_dodge(width=0.9))+  
 geom\_col(aes(y=`5-9 kat\_arasi`, fill="5-9 kat\_arasi"), position=position\_dodge(width=0.9))+  
 geom\_col(aes(y=`9-19 kat\_arasi`, fill="9-19 kat\_arasi"), position=position\_dodge(width=0.9))+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe")+  
 ylab("Bina Sayısı")+  
 labs(title="İlçe Bazında Kat Sayısına Göre Bina Sayıları")  
p14



**“riskli\_yapi\_durumu” analysis:**

An analysis of individual perceptions regarding the structural risk of their households have been conducted.

“İlçe bazında hanelerdeki riskli yapı durumu hakkındaki bireysel öngörülerin oranları” (3 types of individual foresights):

It has been observed that a significant number of people answered “yes” or “I don’t know.”.

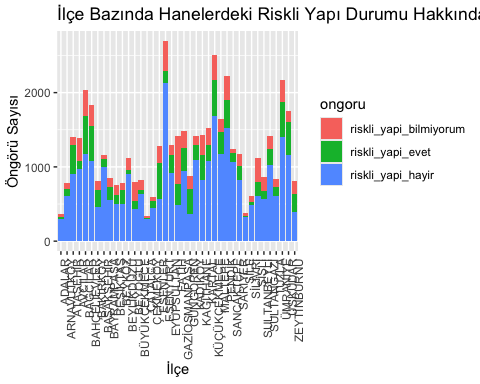
evet<-sum(riskli\_yapi\_durumu$riskli\_yapi\_evet)  
hayir<-sum(riskli\_yapi\_durumu$riskli\_yapi\_hayir)  
bilmiyorum<-sum(riskli\_yapi\_durumu$riskli\_yapi\_bilmiyorum)  
oranlar3<-c(riskli\_yapi\_evet=evet/(evet+hayir+bilmiyorum),   
 riskli\_yapi\_hayir=hayir/(evet+hayir+bilmiyorum),   
 riskli\_yapi\_bilmiyorum=bilmiyorum/(evet+hayir+bilmiyorum))  
print(oranlar3)

riskli\_yapi\_evet riskli\_yapi\_hayir riskli\_yapi\_bilmiyorum   
 0.1793526 0.6675207 0.1531267

“İlçe bazında hanelerdeki riskli yapı durumu hakkındaki bireysel öngörülerin dağılımı” (3 types of individual foresights):

In some districts, the green segments representing those who believe they live in risky structures are particularly noticeable.

#veri çerçevesini uzun formata dönüştürme  
library(tidyr)  
riskli\_yapi\_durumu\_duzenlenmis<-pivot\_longer(riskli\_yapi\_durumu, cols=starts\_with("riskli\_yapi\_"), names\_to="ongoru", values\_to="deger")  
  
p15<-ggplot(riskli\_yapi\_durumu\_duzenlenmis, aes(x=ilce, y=deger, fill=ongoru))+  
 geom\_bar(stat="identity", position="stack")+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe") +  
 ylab("Öngörü Sayısı") +  
 labs(title="İlçe Bazında Hanelerdeki Riskli Yapı Durumu Hakkındaki Bireysel Öngörüler")  
p15



**“kentsel\_donusum\_fikri” analysis:**

An analysis of individual perceptions regarding the idea of urban transformation of their households have been conducted.

“İlçe bazında hanelerdeki kentsel dönüşüm fikri hakkındaki bireysel öngörülerin oranları” (2 types of individual foresights):

It has been observed that there is a significant number of people who consider urban transformation necessary and cannot be underestimated.

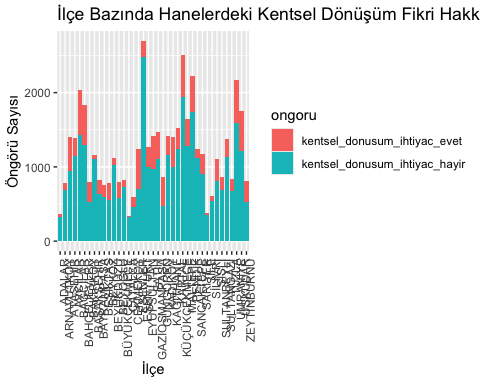
evet<-sum(kentsel\_donusum\_fikri$kentsel\_donusum\_ihtiyac\_evet)  
hayir<-sum(kentsel\_donusum\_fikri$kentsel\_donusum\_ihtiyac\_hayir)  
oranlar4<-c(kentsel\_donusum\_ihtiyac\_evet=evet/(evet+hayir),   
 kentsel\_donusum\_ihtiyac\_hayir=hayir/(evet+hayir))  
print(oranlar4)

kentsel\_donusum\_ihtiyac\_evet kentsel\_donusum\_ihtiyac\_hayir   
 0.2255722 0.7744278

“İlçe bazında hanelerdeki kentsel dönüşüm fikri hakkındaki bireysel öngörülerin dağılımı” (2 types of individual foresights):

In some districts, the red segments representing those who believe in the need of urban transformation are particularly noticeable.

#veri çerçevesini uzun formata dönüştürme  
library(tidyr)  
kentsel\_donusum\_fikri\_duzenlenmis<-pivot\_longer(kentsel\_donusum\_fikri, cols=starts\_with("kentsel\_donusum\_ihtiyac\_"), names\_to="ongoru", values\_to="deger")  
  
p16<-ggplot(kentsel\_donusum\_fikri\_duzenlenmis, aes(x=ilce, y=deger, fill=ongoru))+  
 geom\_bar(stat="identity", position="stack")+  
 theme(axis.text.x=element\_text(angle=90, hjust=1))+  
 xlab("İlçe") +  
 ylab("Öngörü Sayısı") +  
 labs(title="İlçe Bazında Hanelerdeki Kentsel Dönüşüm Fikri Hakkındaki Bireysel Öngörüler")  
p16

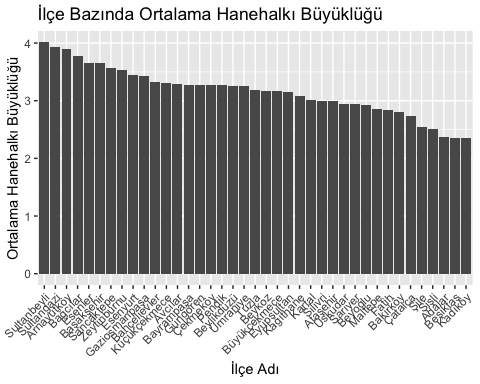


**“ort\_hane” analysis:**

“İlçe Bazlı Ortalama Hanehalkı Büyüklüğü”:

In the analysis of the “ort\_hane” dataset, districts have been ranked from highest to lowest average household size and it has been observed that the most crowded households are in Sultanbeyli.

#ortalama hanehalkı büyüklüğüne göre büyükten küçüğe sıralama  
ort\_hane2<-ort\_hane%>%mutate(ilce=fct\_reorder(ilce, ortalama\_hanehalki\_buyuklugu, .desc=TRUE))  
  
p17<-ggplot(ort\_hane2, aes(x=ilce, y=ortalama\_hanehalki\_buyuklugu))+  
 geom\_bar(stat="identity")+  
 theme(axis.text.x=element\_text(angle=45, hjust=1))+  
 xlab("İlçe Adı")+  
 ylab("Ortalama Hanehalkı Büyüklüğü")+  
 labs(title="İlçe Bazında Ortalama Hanehalkı Büyüklüğü")  
p17



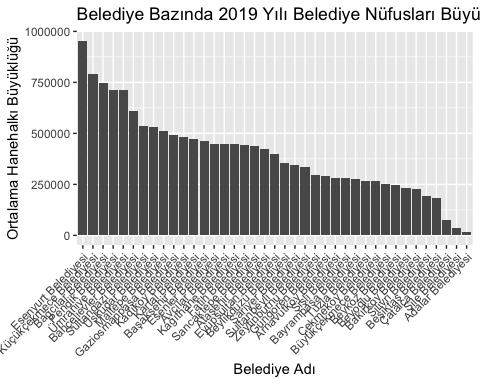
We can see from the graph that in more than half of the districts in İstanbul, the average household size is more than 3.

**“belediye\_nufus” analysis:**

“2019 Yılı Belediye Nüfusları” Grafiği:

In the analysis of the “belediye\_nufus” dataset, municipalities have been ranked from highest to lowest population and it has been observed that the most crowded municipality is the Esenyurt Municipality.

#ortalama hanehalkı büyüklüğüne göre büyükten küçüğe sıralama  
belediye\_nufus2<-belediye\_nufus%>%mutate(belediye=fct\_reorder(belediye, nufus\_2019, .desc=TRUE))  
  
p18<-ggplot(belediye\_nufus2, aes(x=belediye, y=nufus\_2019))+  
 geom\_bar(stat="identity")+  
 theme(axis.text.x=element\_text(angle=45, hjust=1))+  
 xlab("Belediye Adı")+  
 ylab("Ortalama Hanehalkı Büyüklüğü")+  
 labs(title="Belediye Bazında 2019 Yılı Belediye Nüfusları Büyüklüğü")  
p18



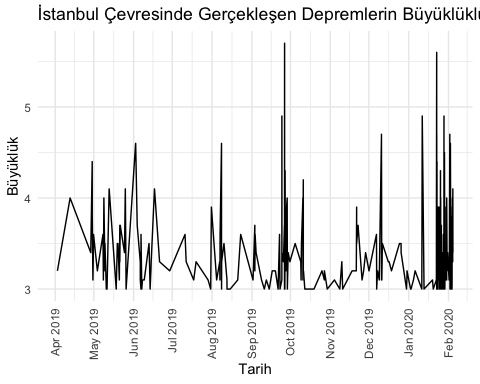
We can visualize the dataset titled “2019 Yılı Belediye Nüfusları”. We can see from the graph that 2 municipalities, Esenyurt and Küçükçekmece, are over 750.000. We can also see from the graph that the first one-third of the municipalities are over 500.000. From this, we can identify the most populous municipalities that would need the most help in an earthquake disaster.

**“gerceklesen\_deprem” analysis:**

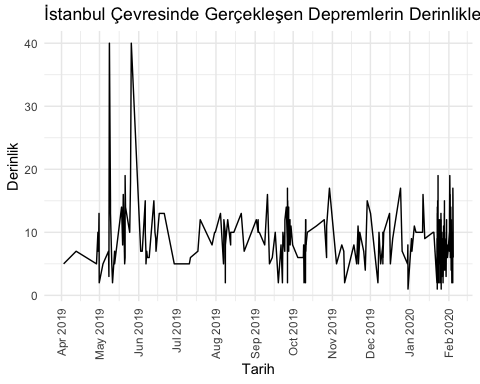
Analyzing this data could be crucial for examining earthquake activity and identifying potential risks. However, predicting future earthquake magnitudes is a complex thing to do because earthquakes are random and unpredictible.

“İstanbul Çevresinde Gerçekleşen Depremlerin Büyüklüklerinin Zaman Serisi Grafiği”:

#time sütununu işleme  
gerceklesen\_deprem$time<-as.POSIXct(gerceklesen\_deprem$time, format="%Y-%m-%dT%H:%M:%OSZ")  
  
library(ggplot2)  
  
p19<-ggplot(gerceklesen\_deprem, aes(x=time, y=magnitude))+  
 geom\_line()+  
 labs(title="İstanbul Çevresinde Gerçekleşen Depremlerin Büyüklüklükleri", x="Tarih", y="Büyüklük")+  
 theme\_minimal()+  
 scale\_x\_datetime(date\_breaks="1 month", date\_labels="%b %Y")+  
 theme(axis.text.x=element\_text(angle=90, vjust=0.5, hjust=1))  
p19



p20<-ggplot(gerceklesen\_deprem, aes(x=time, y=`depth/km`))+  
 geom\_line()+  
 labs(title="İstanbul Çevresinde Gerçekleşen Depremlerin Derinlikleri", x="Tarih", y="Derinlik")+  
 theme\_minimal()+  
 scale\_x\_datetime(date\_breaks="1 month", date\_labels="%b %Y")+  
 theme(axis.text.x=element\_text(angle=90, vjust=0.5, hjust=1))  
p20



What stands out in both graphs is the increase in earthquake frequencies after January 2020. This indicates that the danger is coming.

“İstanbul Çevresinde Gerçekleşen Depremlerin Harita Üzerinde Gösterimi”:

The created map allows us to observe whre earthquakes are concentrated.

#leaflet paketini yükleme  
#install.packages("leaflet")  
library(leaflet)  
  
#veri çerçevesi oluşturma  
depremler<-data.frame(name=gerceklesen\_deprem$event\_id,  
 latitude=gerceklesen\_deprem$latitude,  
 longitude=gerceklesen\_deprem$longitude)  
  
#İstanbul haritasını oluşturma  
istanbul\_map<-leaflet()%>%setView(lng=28.9784, lat=41.0082, zoom=11)%>%addTiles()   
  
#verileri haritaya ekleme  
for(i in 1:nrow(depremler)){  
 istanbul\_map<-addMarkers(istanbul\_map, lng=depremler$longitude[i], lat=depremler$latitude[i], popup=depremler$name[i])  
 }  
  
#haritayı görüntüleme  
istanbul\_map

### District Based Overall Score Calculation, Sorting and Risk Analysis

A dataframe named “analiz” was constructed in order to contain the data to be used for calculating the overall risk score on a district basis. The rows represent the districts and the columns represent the factors. Each factor has a weight in order to calculate the overall risk score and the weights are located in the “weight” vector.

#mahalle bazında veri içeren verisetlerinde ilçe bazında olacak şekilde düzeltmeler yapma  
deprem\_senaryo2<-deprem\_senaryo%>%group\_by(ilce\_adi)%>%summarise(across(starts\_with("cok\_agir\_hasarli\_bina\_sayisi"):starts\_with("gecici\_barinma"), sum))  
bina\_sayi2<-bina\_sayi%>%group\_by(ilce\_adi)%>%summarise(across(starts\_with("1980\_oncesi"):starts\_with("9-19 kat\_arasi"), sum))  
  
#"analiz" adlı bir dataframe oluşturma  
analiz<-data.frame(ilce\_adi=deprem\_senaryo2$ilce\_adi,  
 cok\_agir\_hasarli\_bina\_sayisi=deprem\_senaryo2$cok\_agir\_hasarli\_bina\_sayisi,  
 agir\_hasarli\_bina\_sayisi=deprem\_senaryo2$agir\_hasarli\_bina\_sayisi,  
 orta\_hasarli\_bina\_sayisi=deprem\_senaryo2$orta\_hasarli\_bina\_sayisi,  
 hafif\_hasarli\_bina\_sayisi=deprem\_senaryo2$hafif\_hasarli\_bina\_sayisi,  
 can\_kaybi\_sayisi=deprem\_senaryo2$can\_kaybi\_sayisi,  
 agir\_yarali\_sayisi=deprem\_senaryo2$agir\_yarali\_sayisi,  
 hastanede\_tedavi\_sayisi=deprem\_senaryo2$hastanede\_tedavi\_sayisi,  
 hafif\_yarali\_sayisi=deprem\_senaryo2$hafif\_yarali\_sayisi,  
 dogalgaz\_boru\_hasari=deprem\_senaryo2$dogalgaz\_boru\_hasari,  
 icme\_suyu\_boru\_hasari=deprem\_senaryo2$icme\_suyu\_boru\_hasari,  
 atik\_su\_boru\_hasari=deprem\_senaryo2$atik\_su\_boru\_hasari,  
 gecici\_barinma=deprem\_senaryo2$gecici\_barinma,  
 `1980\_oncesi`=bina\_sayi2$`1980\_oncesi`,  
 `1980-2000\_arasi`=bina\_sayi2$`1980-2000\_arasi`,  
 `2000\_sonrasi`=bina\_sayi2$`2000\_sonrasi`,  
 `1-4 kat\_arasi`=bina\_sayi2$`1-4 kat\_arasi`,  
 `5-9 kat\_arasi`=bina\_sayi2$`5-9 kat\_arasi`,  
 `9-19 kat\_arasi`=bina\_sayi2$`9-19 kat\_arasi`,  
 ortalama\_hanehalki\_buyuklugu=ort\_hane$ortalama\_hanehalki\_buyuklugu)  
  
#"weight" adlı bir vektör oluşturma  
weight<-c(0.09,0.07,0.05,0.03,0.09,0.07,0.05,0.03,0.05,0.05,0.05,0.05,0.08,0.05,0.03,0.03,0.05,0.08)  
  
#overall score hesaplama  
overall\_risk\_score<-analiz$cok\_agir\_hasarli\_bina\_sayisi\*weight[1]+  
 analiz$agir\_hasarli\_bina\_sayisi\*weight[2]+  
 analiz$orta\_hasarli\_bina\_sayisi\*weight[3]+  
 analiz$hafif\_hasarli\_bina\_sayisi\*weight[4]+  
 analiz$can\_kaybi\_sayisi\*weight[5]+  
 analiz$agir\_yarali\_sayisi\*weight[6]+  
 analiz$hastanede\_tedavi\_sayisi\*weight[7]+  
 analiz$hafif\_yarali\_sayisi\*weight[8]+  
 analiz$dogalgaz\_boru\_hasari\*weight[9]+  
 analiz$icme\_suyu\_boru\_hasari\*weight[10]+  
 analiz$atik\_su\_boru\_hasari\*weight[11]+  
 analiz$gecici\_barinma\*weight[12]+  
 analiz$X1980\_oncesi\*analiz$ortalama\_hanehalki\_buyuklugu\*weight[13]+  
 analiz$X1980.2000\_arasi\*analiz$ortalama\_hanehalki\_buyuklugu\*weight[14]+  
 analiz$X2000\_sonrasi\*analiz$ortalama\_hanehalki\_buyuklugu\*weight[15]+  
 analiz$X1.4.kat\_arasi\*analiz$ortalama\_hanehalki\_buyuklugu\*weight[16]+  
 analiz$X5.9.kat\_arasi\*analiz$ortalama\_hanehalki\_buyuklugu\*weight[17]+  
 analiz$X9.19.kat\_arasi\*analiz$ortalama\_hanehalki\_buyuklugu\*weight[18]  
  
latitude=c(40.8747,41.1864,40.9833,40.9792,41.0341,40.9977,40.9804,41.0837,41.0349,41.0441,41.1271,41.0133,41.0371,41.0248,41.1421,41.0323,41.0542,41.0412,41.0551,41.0203,41.0576,41.0105,40.9903,41.0717,40.8999,41.0092,40.9339,40.8796,41.0090,41.1664,41.1749,41.0737,41.0604,40.9684,41.1070,40.8144,41.0338,41.0327,40.9910)  
longitude=c(29.1294,28.7389,29.1278,28.7214,28.8330,28.8506,28.8724,28.8169,28.9122,29.0017,29.0978,28.6489,28.9774,28.5854,28.4575,29.1695,28.8676,28.6939,28.9346,28.9339,28.9153,28.8741,29.0205,28.9646,29.1936,28.7757,29.1650,29.2580,29.2109,29.0500,29.6096,28.2479,28.9878,29.2620,28.8714,29.3094,29.1013,29.0319,28.8968)  
  
overall\_risk\_score\_df<-data.frame(ilce\_adi=analiz$ilce\_adi, latitude=latitude, longitude=longitude, overall\_risk\_score=overall\_risk\_score)

After the overall score calculation was completed, ranking was performed and the districts were ranked from highest risk to lowest risk. Finally, the overall scores were shown on the map.

overall\_risk\_score\_df<-overall\_risk\_score\_df[order(overall\_risk\_score\_df$overall\_risk\_score, decreasing=TRUE), ]  
overall\_risk\_score\_df

ilce\_adi latitude longitude overall\_risk\_score  
18 FATİH 41.0412 28.6939 20835.981  
8 BAĞCILAR 41.0837 28.8169 18841.300  
26 PENDİK 41.0092 28.7757 16853.602  
24 KÜÇÜKÇEKMECE 41.0717 28.9646 16487.425  
16 ESENYURT 41.0323 29.1695 15474.934  
36 ÜMRANİYE 40.8144 29.3094 15326.482  
37 ÜSKÜDAR 41.0338 29.1013 14236.909  
28 SARIYER 40.8796 29.2580 14050.153  
5 BAHÇELİEVLER 41.0341 28.8330 13016.060  
17 EYÜP 41.0542 28.8676 12666.458  
31 SİLİVRİ 41.1749 29.6096 10846.386  
10 BEYKOZ 41.0441 29.0017 10844.370  
22 KARTAL 41.0105 28.8741 10801.665  
29 SULTANBEYLİ 41.0090 29.2109 10595.132  
12 BEYOĞLU 41.0133 28.6489 10499.915  
21 KADIKÖY 41.0576 28.9153 10492.666  
19 GAZİOSMANPAŞA 41.0551 28.9346 10119.521  
4 AVCILAR 40.9792 28.7214 10001.234  
2 ARNAVUTKÖY 41.1864 28.7389 9778.344  
25 MALTEPE 40.8999 29.1936 9714.603  
30 SULTANGAZİ 41.1664 29.0500 9625.334  
14 BÜYÜKÇEKMECE 41.0248 28.5854 9418.842  
15 ESENLER 41.1421 28.4575 8572.687  
3 ATAŞEHİR 40.9833 29.1278 8534.802  
39 ŞİŞLİ 40.9910 28.8968 8420.029  
32 TUZLA 41.0737 28.2479 8395.400  
34 ÇATALCA 40.9684 29.2620 8070.917  
27 SANCAKTEPE 40.9339 29.1650 7792.004  
7 BAYRAMPAŞA 40.9804 28.8724 7772.034  
9 BAŞAKŞEHİR 41.0349 28.9122 7732.653  
23 KAĞITHANE 40.9903 29.0205 7608.464  
35 ÇEKMEKÖY 41.1070 28.8714 6403.390  
6 BAKIRKÖY 40.9977 28.8506 6223.436  
33 ZEYTİNBURNU 41.0604 28.9878 6084.945  
11 BEYLİKDÜZÜ 41.1271 29.0978 5217.615  
20 GÜNGÖREN 41.0203 28.9339 5144.587  
13 BEŞİKTAŞ 41.0371 28.9774 4984.912  
38 ŞİLE 41.0327 29.0319 4655.740  
1 ADALAR 40.8747 29.1294 1856.583

#en riskli ilçeyi belirleme  
highest\_overall\_risk\_score<-overall\_risk\_score\_df[which.max(overall\_risk\_score\_df$overall\_risk\_score), "ilce\_adi"]  
  
risk\_map<-leaflet(overall\_risk\_score\_df)%>%addTiles()%>%addCircleMarkers(lng=~overall\_risk\_score\_df$longitude, lat=~overall\_risk\_score\_df$latitude, label=~paste0(overall\_risk\_score\_df$ilce\_adi, "<br>", "Overall Skor: ", overall\_risk\_score\_df$overall\_risk\_score), color=ifelse(overall\_risk\_score\_df$ilce\_adi==highest\_overall\_risk\_score, "red", "blue"), radius=5, stroke=FALSE, fillOpacity=0.8)%>%addLegend(position="bottomright", colors=c("red", "blue"), labels=c("En Riskli İlçe", "Diğer İlçeler"), title="Risk Seviyeleri")  
risk\_map

## 3.2 Results

As a result of these analyses, districts where precautions need to be taken, where the most loss of life and property could occur, and where potential interventions need to be well planned have been identified. These are Fatih, Bağcılar, Pendik, Küçükçekmece and Esenyurt. If the weights used in calculating the overall risk score are determined by experts, much more reliable and accurate results will be obtained.

# 4. Results and Key Takeaways 📈

Istanbul is Europe’s most populous city. It means that the potential of destruction and the need for help would be very high if a disaster such as an earthquake occurs.

The purpose of this project is to examine the risky structural condition, to make district-based interpretations and to evaluate the opinions of people living in Istanbul.

7 datasets were used in this study, plots were drawn and interpreted comprehensively.

* This study has determined which districts are most likely to experience possible building and infrastructure damage in an earthquake disaster. This helps determine where, how and how much aid should be directed.
* By calculating the overall risk scores, the districts of Istanbul have been ranked from the most risky to least risky. Thanks to his ranking, the government will be able to take appropriate steps both before and after an earthquake, aiding in preparedness and response efforts.
* Lastly, this study serves as both a guide and an encouragement for the society to be prepared for this disaster.

This study has

* identified the regions most prone to earthquakes and the cities most at risk,
* depicted them on maps,
* drawn conclusions about the opinions of survey participants and
* examined building and infrastructure in Istanbul.

\*ChatGPT was utilized in this project.