

PYTHON

Bu kodlar fayl şəklində göndəriləcək. Ətraflı şəkildə oradan baxa bilərsiniz.
Birinci olaraq verilmiş suallara baxaq.

1. Ümumi müraciət sayı

```
data['Müraciət edən vətəndaşın eyniləşdirmə nömrəsi'].count()
```

50000

2. Müraciət edən unikal vətəndaş sayı

```
data['Müraciət edən vətəndaşın eyniləşdirmə nömrəsi'].nunique()
```

44103

3. Müraciətlərin cins üzrə paylanması(0 - Qadın, 1 - Kişi)

```
data.groupby('Cins')['Cins'].count().sort_values( axis=0, ascending=False)
```

```
Cins
1    36075
0    13925
Name: Cins, dtype: int64
```

```
man_per= (data[data['Cins']==1].shape[0] *100)/(data['Cins'].shape[0])
man_per
```

72.15

```
woman_per= (data[data['Cins']==0].shape[0] *100)/(data['Cins'].shape[0])
woman_per
```

27.85

4. Ən çox müraciət edilən 5 xidməti

```
data.groupby('Xidmət')['Xidmət'].count().sort_values( axis=0, ascending=False).head(5)
```

```
Xidmət
Haqqı ödənilən ictimai işlər                3964
Özünüməşğulluq (kənd təsərrüfatı zərfləri üzrə )  3936
Peşə yönümlü məsləhət                      3905
Əməkhaqqının birgə maliyyələşdirilməsi      3893
Təcrübə proqramlarının təşkili             3888
Name: Xidmət, dtype: int64
```

5. Təhsil səviyyələri üzrə müraciət sayları

Orta təhsilli şəxslər daha çox müraciət etmişdir.

```
data.groupby('Təhsil səviyyəsi')['Təhsil səviyyəsi'].count().sort_values( axis=0, ascending=False)
```

```
Təhsil səviyyəsi
Ümumi təhsil-tam orta təhsil      6744
Ali təhsil-bakalavriat           4041
Ümumi təhsil-ümumi orta təhsil    1888
Tam orta təhsil bazasında orta ixtisas təhsili  1560
Orta ixtisas təhsili              1556
Ümumi orta təhsil bazasında orta ixtisas təhsili  869
İlk peşə-ixtisas təhsili         777
Ali təhsil-bakalavriat, Ali təhsil-magistratura  321
bakalavriat                      302
Ümumi təhsil-ibtidai təhsil       27
Name: Təhsil səviyyəsi, dtype: int64
```

6. Ən az müraciət edilən 3 filialı tapın

```
data.groupby('Müraciət edilmiş filial')['Müraciət edilmiş filial'].count().sort_values( axis=0, ascending=False).tail(3)
```

```
Müraciət edilmiş filial
Bakı İqtisadi rayonu      9
Şərqi Zəngəzur İR        9
Mil-Muğan İqtisadi rayonu 5
Name: Müraciət edilmiş filial, dtype: int64
```

1. Kitabxanaları daxil edirik. Verilmiş datanı üzərində işləmək üçün kodumuza əlavə edirik.

Və ilk 5 sətirinin necə görünməsinə baxırıq.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as seabornInstance
%matplotlib inline
```

```
data = pd.read_excel('data.xlsx', sheet_name='Data')
```

```
data.head()
```

	Müraciətin eynileşdirmə nömrəsi	Müraciət tarixi	Müraciət edən vətəndaşın eynileşdirmə nömrəsi	Doğum tarixi	Cins	Ailə vəziyyəti	Xidmət	Müraciət edilmiş filial	Təhsil səviyyəsi	Faktiki ünvan
0	1002307	2019-03-11 00:00:00	10012	1978-12- 07	1	Evli	İşsizlikdən sığorta ödənişin təşkili	Qazax-Tovuz	Ali təhsil- bakalavriat	Şamaxı rayonu, Keşlə qəs. M.ARIF pr., b. 9, m. 8
1	15317345	2020-04-25 00:00:00	10037	1995-01- 18	0	Evli	İşsizlikdən sığorta ödənişin təşkili	Naxçıvan	bakalavriat	Qubadlı rayonu, MƏMMƏD ARAZ küç., b. 1, m. 157
2	1003198	2019-03-11 00:00:00	10084	1992-08- 27	1	Evli	Özünüməşğulluq (kənd təsərrüfatı zərfləri üzrə)	Mil-Muğan İqtisadi rayonu	bakalavriat	Mingəçevir ş., RƏHİM ŞİXƏLİYEV pr., b. 8, m. 08
3	1002647	2019-03-11 00:00:00	10148	1991-01- 30	0	Evli	Peşə hazırlığı xidmətləri	Gəncə- Daşkəsən	bakalavriat	Gəncə ş., XURŞUD AĞAYEV prospekti, b. 5, m. 139
4	1002701	2019-03-11 00:00:00	10150	2000-01- 24	0	Evli	Peşə hazırlığı xidmətləri	Mərkəzi Aran	bakalavriat	Kürdəmir rayonu, ASİF MƏMMƏDOV küçəsi, b. 8, m...

2. Verilmiş sütunların tiplərinə , neçəsinin dolu olmasına baxırıq.

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50000 entries, 0 to 49999
Data columns (total 10 columns):
 #   Column                                                                 Non-Null Count  Dtype
---  -
 0   Müraciətin eyniləşdirmə nömrəsi                                     50000 non-null  int64
 1   Müraciət tarixi                                                    49998 non-null  object
 2   Müraciət edən vətəndaşın eyniləşdirmə nömrəsi                     50000 non-null  int64
 3   Doğum tarixi                                                       49993 non-null  datetime64[ns]
 4   Cins                                                                50000 non-null  int64
 5   Ailə vəziyyəti                                                    49992 non-null  object
 6   Xidmət                                                            50000 non-null  object
 7   Müraciət edilmiş filial                                           50000 non-null  object
 8   Təhsil səviyyəsi                                                  18085 non-null  object
 9   Faktiki ünvan                                                      50000 non-null  object
dtypes: datetime64[ns](1), int64(3), object(6)
memory usage: 3.8+ MB
```

3. Hər bir sütunda cəmi neçə boş sətir olduğuna baxaq

```
data.isnull().sum()
```

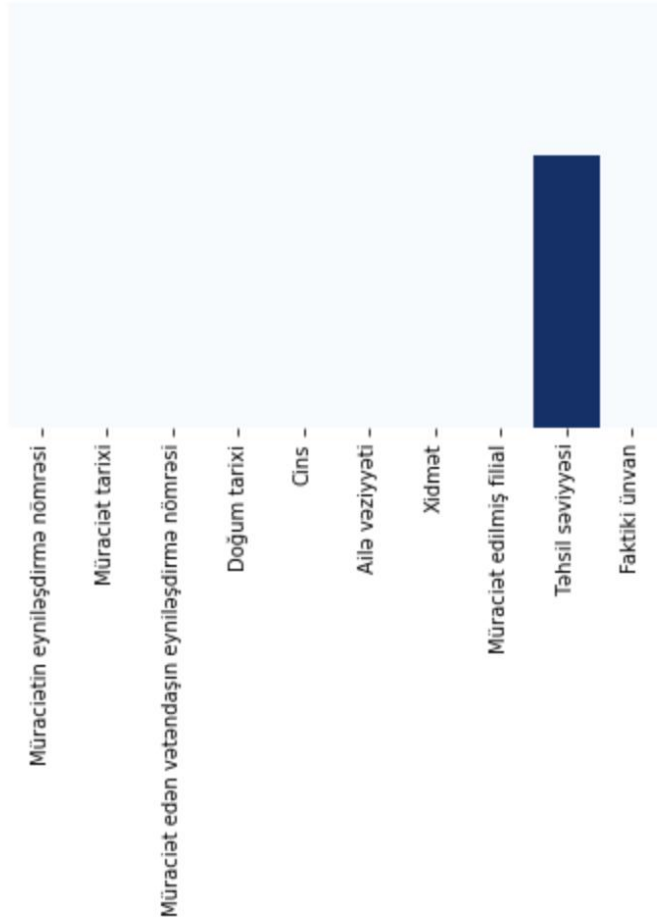
```
Müraciətin eyniləşdirmə nömrəsi      0
Müraciət tarixi                      2
Müraciət edən vətəndaşın eyniləşdirmə nömrəsi  0
Doğum tarixi                         7
Cins                                  0
Ailə vəziyyəti                       8
Xidmət                               0
Müraciət edilmiş filial              0
Təhsil səviyyəsi                    31915
Faktiki ünvan                        0
dtype: int64
```

Görüldüyü kimi “Müraciət tarixi”, “Doğum tarixi” və “Ailə vəziyyətin”-də boş sətirlər az saydadır . Bu bizim nəticəmizə təsir etməyəcəkdir. Lakin “Təhsil səviyyə”-nin 50%-dəndə çoxu boşdur . Buda müraciət edənlərin təhsil səviyyələri haqqında ümumi fikrə gəlməyimizə mane olur .

4. Qrafik vasitəsi ilə də aydın şəkildə boş sətirlərin sayını görə bil

```
import seaborn as sns
sns.heatmap(data.isnull(),yticklabels=False,cbar=False,cmap='Blues')
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f940e4b6890>



5. Verilmiş doğum tarixləri vasitəsilə müraciət edən şəxslərin yaşlarını hesablayıb və bunu “yaş” adlı sütuna əlavə etdim

```
data['yaş']=pd.to_datetime('today').year-pd.to_datetime(data['Doğum tarixi']).dt.year
# (pd.to_datetime('today').year-pd.to_datetime('1956-07-01').year)
data.head()
```

	Müraciətin eyniləşdirmə nömrəsi	Müraciət tarixi	Müraciət edən vətəndaşın eyniləşdirmə nömrəsi	Doğum tarixi	Cins	Ailə vəziyyəti	Xidmət	Müraciət edilmiş filial	Təhsil səviyyəsi	Faktiki ünvan	yaş
0	1002307	2019-03-11 00:00:00	10012	1978-12-07	1	Evli	İşsizlikdən sığorta ödənişin təşkil	Qazax-Tovuz	Ali təhsil-bakalavriat	Şamaxı rayonu, Keşlə qəs., M.ARIF pr., b. 9, m. 8	44.0
1	15317345	2020-04-25 00:00:00	10037	1995-01-18	0	Evli	İşsizlikdən sığorta ödənişin təşkil	Naxçıvan	bakalavriat	Qubadlı rayonu, MƏMMƏD ARAZ küç., b. 1, m. 157	27.0
2	1003198	2019-03-11 00:00:00	10084	1992-08-27	1	Evli	Özünüməşğulluq (kənd təsərrüfatı zərfləri üzrə)	Mil-Muğan İqtisadi rayonu	bakalavriat	Mingəçevir ş., RƏHİM ŞİXƏLİYEV pr., b. 8, m. 08	30.0
3	1002647	2019-03-11 00:00:00	10148	1991-01-30	0	Evli	Peşə hazırlığı xidmətləri	Gəncə-Daşkəsən	bakalavriat	Gəncə ş., XURŞUD AĞAYEV prospekti, b. 5, m. 139	31.0
4	1002701	2019-03-11 00:00:00	10150	2000-01-24	0	Evli	Peşə hazırlığı xidmətləri	Mərkəzi Aran	bakalavriat	Kürdəmir rayonu, ASİF MƏMMƏDOV küçəsi, b. 8, m...	22.0

6. Müraciət edənlərin ortalama yaşı 32-dir.

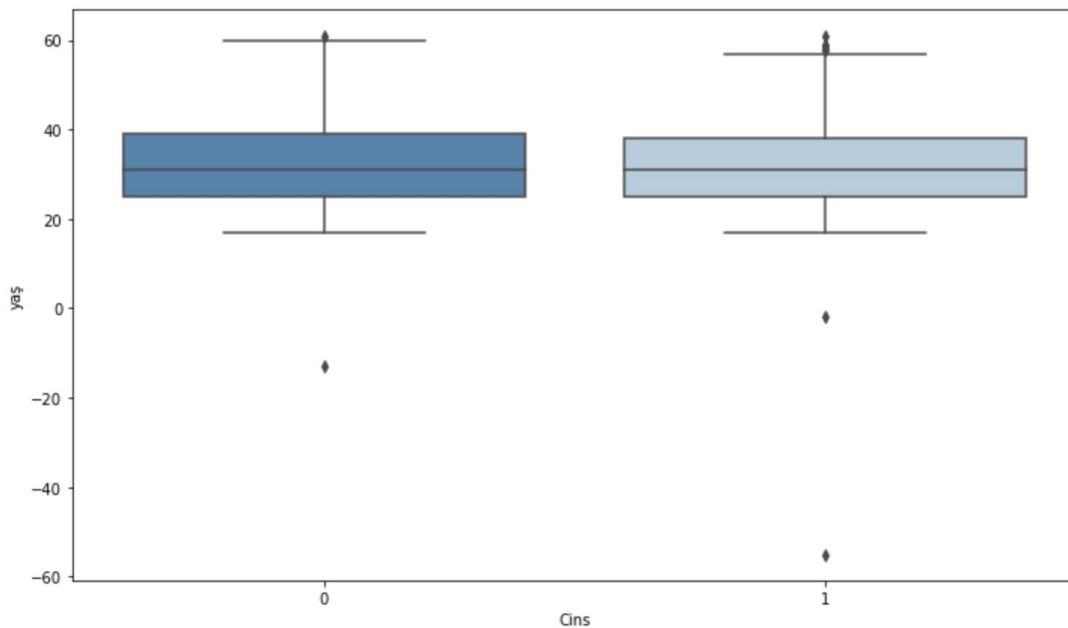
```
data['yaş'].mean()
```

32.05952833396675

7. Datada çirklə məlumatlar var . Qrafikdən görüldüyü kimi yaşı 0-dan aşağı olan sətirlərimiz var.

```
plt.figure(figsize=(12, 7))
sns.boxplot(x='Cins',y='yaş',data=data,palette='Blues_r')
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f940e301650>



8. Yaşları boş olan sətirləri ortalama yaş ilə əvəz edirəm

```
data['yaş'].fillna(value=data['yaş'].mean(), inplace=True)
```

9. Olan datalarda ən çox istifadə olunan sətir

```
data['Təhsil səviyyəsi'].mode()
```

0 Ümumi təhsil-tam orta təhsil
dtype: object

10. “Təhsil səviyyəsi” sütununda boş datalar həddindən çox olduğundan o sütunu silirik.

```
data.drop('Təhsil səviyyəsi',axis=1,inplace=True)
data.dropna(inplace=True)
data.head()
```

	Müraciətin eyniləşdirmə nömrəsi	Müraciət tarixi	Müraciət edən vətəndaşın eyniləşdirmə nömrəsi	Doğum tarixi	Cins	Ailə vəziyyəti	Xidmət	Müraciət edilmiş filial	Faktiki ünvan	yaş
0	1002307	2019-03-11 00:00:00	10012	1978-12- 07	1	Evli	İşsizlikdən sığorta ödənişin təşkili	Qazax-Tovuz	Şamaxı rayonu, Keşlə qəs., M.ARİF pr., b. 9, m. 8	44.0
1	15317345	2020-04-25 00:00:00	10037	1995-01- 18	0	Evli	İşsizlikdən sığorta ödənişin təşkili	Naxçıvan	Qubadlı rayonu, MƏMMƏD ARAZ küç., b. 1, m. 157	27.0
2	1003198	2019-03-11 00:00:00	10084	1992-08- 27	1	Evli	Özünüməşğulluq (kənd təsərrüfatı zərfləri üzrə)	Mil-Muğan İqtisadi rayonu	Mingəçevir ş., RƏHİM ŞİXƏLİYEV pr., b. 8, m. 08	30.0
3	1002647	2019-03-11 00:00:00	10148	1991-01- 30	0	Evli	Peşə hazırlığı xidmətləri	Gəncə-Daşkəsən	Gəncə ş., XURŞUD AĞAYEV prospekti, b. 5, m. 139	31.0
4	1002701	2019-03-11 00:00:00	10150	2000-01- 24	0	Evli	Peşə hazırlığı xidmətləri	Mərkəzi Aran	Kürdəmir rayonu, ASİF MƏMMƏDOV küçəsi, b. 8, m...	22.0

11.Yaşları mənfı olan şəxslərin dəyərini ortalama yaş ilə əvəz etdim

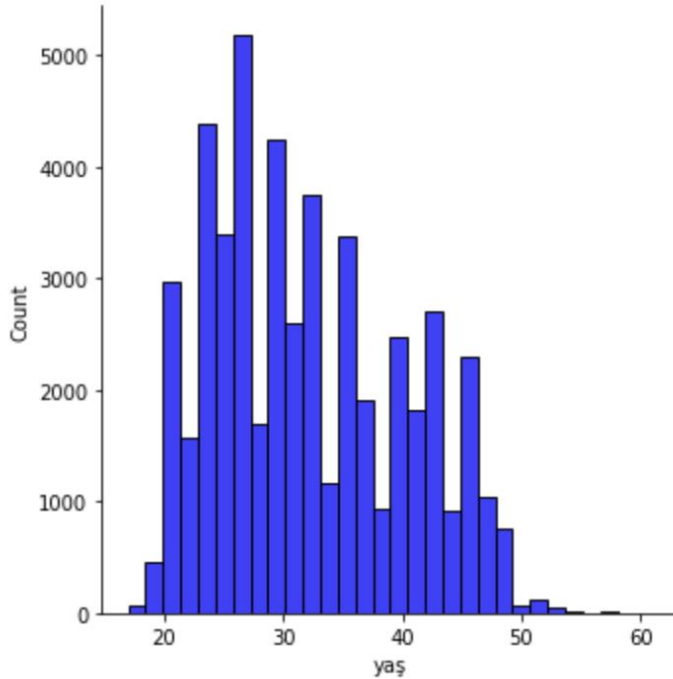
```
] data[data['yaş']<0]=-32
data[data['yaş']<0]
```

Müraciətin eyniləşdirmə nömrəsi	Müraciət tarixi	Müraciət edən vətəndaşın eyniləşdirmə nömrəsi	Doğum tarixi	Cins	Ailə vəziyyəti	Xidmət	Müraciət edilmiş filial	Faktiki ünvan	yaş
---------------------------------	-----------------	---	--------------	------	----------------	--------	-------------------------	---------------	-----

12.Müraciət etmə yaşlarına baxacaq olarsaq ,bunun 25-33 yaş aralıqlarında müraciətlər daha çoxdur . Lakin 48+ olduğu zaman azalır.

```
#distribution plot of age of the people
sns.displot(data['yaş'].dropna(), kde=False, bins=30, color='Blue')
```

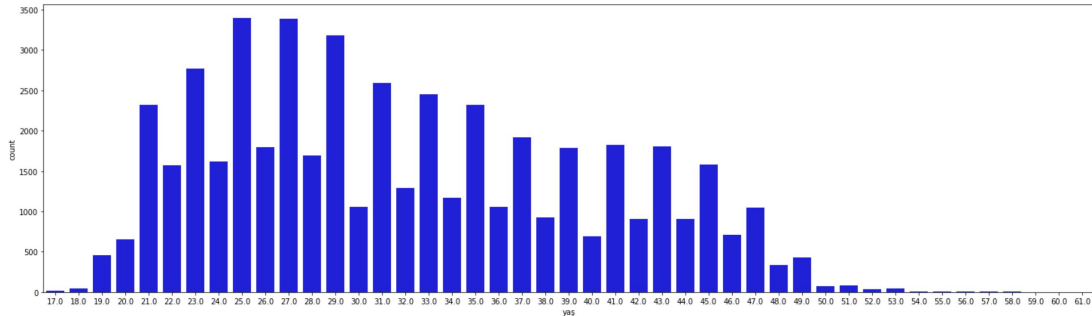
<seaborn.axisgrid.FacetGrid at 0x7f940e7ea910>



13. Bu qrafik vasitəsilə daha aydın görə bilərik.

```
plt.figure(figsize=(25, 7))
sns.countplot(x=data['yaş'],data=data,color="Blue")
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f940967ef50>



14. Gələcək barəsində fikir demək üçün datamızda string və ya object tipli dəyişənlər olmamalıdır. Buna görə onları integer tiplə əvəz edirəm. Və son halı bu şəkildə olur.

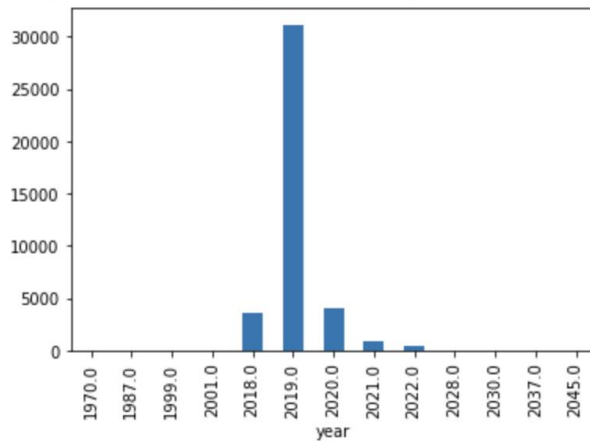
	Müraciətin eyniləşdirmə nömrəsi	Müraciət tarixi	Müraciət edən vətəndaşın eyniləşdirmə nömrəsi	Doğum tarixi	Cins	Ailə vəziyyəti	Xidmət	Müraciət edilmiş filial	Faktiki ünvan	yaş
0	1002307	2019-03-11 00:00:00	10012	1978-12-07 00:00:00	1	3	13	6	39262	44.0
1	15317345	2020-04-25 00:00:00	10037	1995-01-18 00:00:00	0	3	13	4	19715	27.0
2	1003198	2019-03-11 00:00:00	10084	1992-08-27 00:00:00	1	3	12	2	11384	30.0
3	1002647	2019-03-11 00:00:00	10148	1991-01-30 00:00:00	0	3	8	18	3835	31.0
4	1002701	2019-03-11 00:00:00	10150	2000-01-24 00:00:00	0	3	8	3	5604	22.0
...
49995	567298	2019-01-17 00:00:00	54092	1975-04-04 00:00:00	0	4	1	3	22	47.0
49996	569874	2019-01-18 00:00:00	54096	1996-09-08 00:00:00	1	4	5	5	307	26.0
49997	570346	2019-01-31 00:00:00	54097	2001-05-16 00:00:00	1	4	13	5	11042	21.0
49998	21807075	2021-08-03 00:00:00	54097	2001-12-23 00:00:00	1	4	9	7	13144	21.0
49999	570370	2019-01-18 00:00:00	54099	1983-06-05 00:00:00	1	4	11	20	41125	39.0

15. Müraciət tarixi sütununda 3 sətirdə tarix sıralaması düz yazılmamışdı .
(12.17.2019, 02.18.2021, 04.29.2019). Və bu sətirlərin prediction-da problem yaratmaması üçün sildim

16. 2019-cu ildə müraciətlər daha çox olmuşdur . Və qarşıdakı illərdə müraciətlərin daha da azalacağı görünür.

```
train.groupby('year')['Müraciət edən vətəndaşın eyniləşdirmə nömrəsi'].count().plot.bar()
```

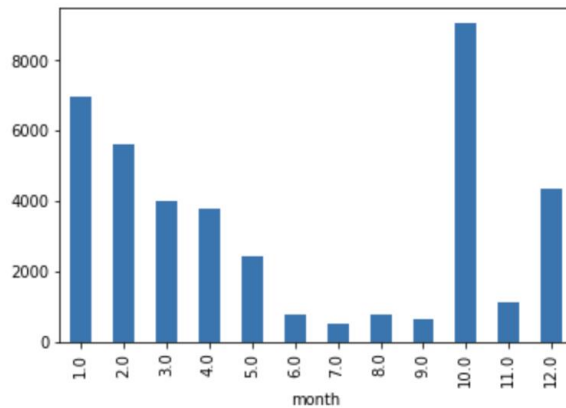
<matplotlib.axes._subplots.AxesSubplot at 0x7f940916f350>



17. Hansı ayda daha çox müraciət olunduğuna baxacaq olarsaq , bu oktyabr ayıdır

```
train.groupby('month')['Müraciət edən vətəndaşın eyniləşdirmə nömrəsi'].count().plot.bar()
```

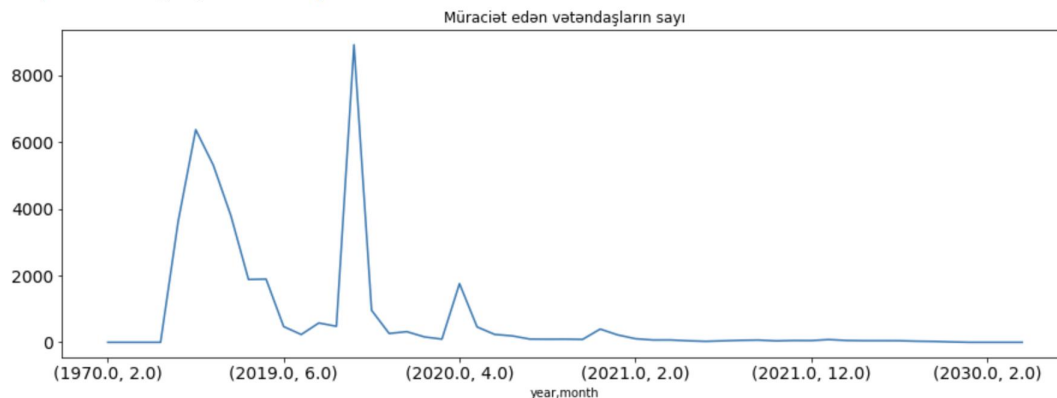
<matplotlib.axes._subplots.AxesSubplot at 0x7f9409a200d0>



18. Müraciət saylarının il və aylar ərzində necə bölündüyünü görürük . Və qarşıdakı illərə doğru say azalır.

```
temp=train.groupby(['year', 'month'])['Müraciət edən vətəndaşın eyniləşdirmə nömrəsi'].count()
temp.plot(figsize=(15,5), title= 'Müraciət edən vətəndaşların sayı', fontsize=14)
```

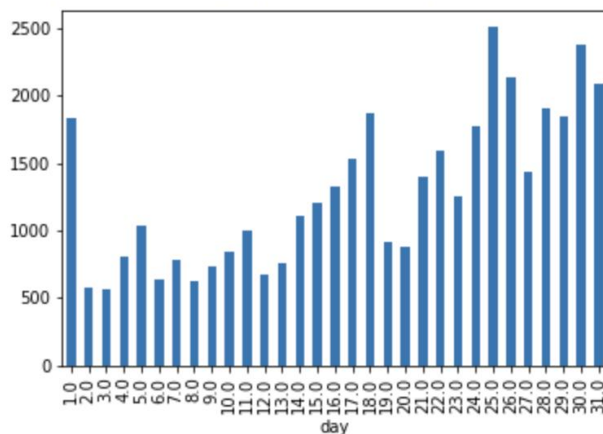
<matplotlib.axes._subplots.AxesSubplot at 0x7f94099ca850>



19. Ayın hansı günlərində daha çox müraciət olunur baxa bilərik. Nəticədə ayın sonlarına doğru müraciət sayları artır.

```
train.groupby('day')['Müraciət edən vətəndaşın eyniləşdirmə nömrəsi'].count().plot.bar()
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f940996fe50>



POWER BI

Bu qrafiklər hazır power bi faylı şəklində atılacaq . İstənilən dəyişikliyi edə və rəqəmlərə baxa bilərsiniz.

Birinci olaraq verilmiş suallara baxaq.

1. Ümumi müraciət sayı

Ümumi müraciət sayı

50000

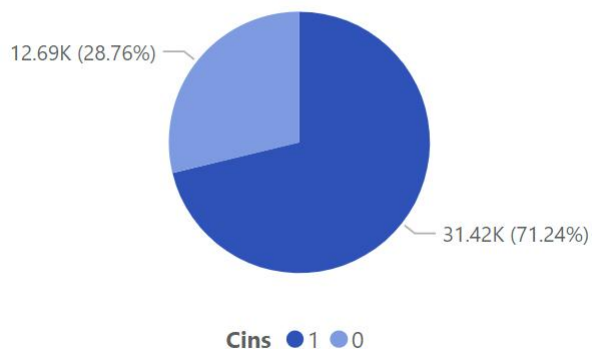
2. Müraciət edən unikal vətəndaş sayı

Müraciət edən unikal vətəndaş sayı

44103

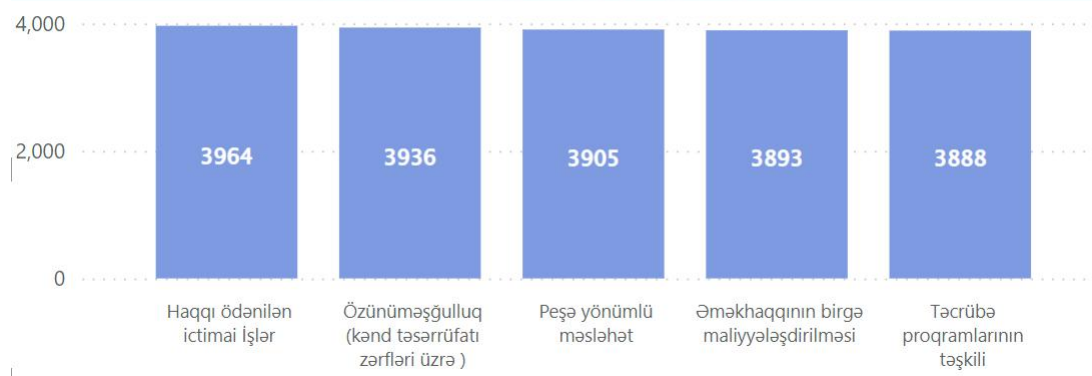
3. Müraciətlərin cins üzrə paylanması(0 - Qadın, 1 - Kişi)

Müraciətlərin cins üzrə paylanması
(0-qadın 1–kişi)



4. Ən çox müraciət edilən 5 xidməti

Ən çox müraciət edilən 5 xidmət

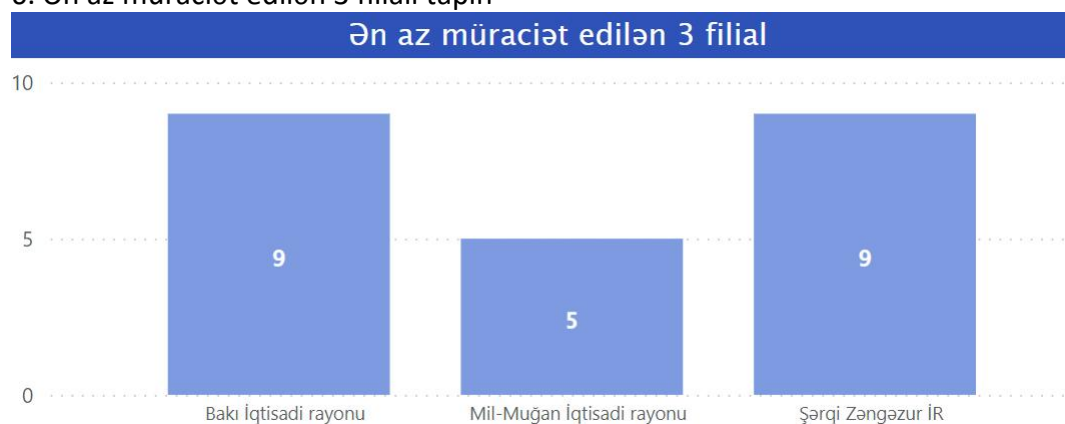


5. Təhsil səviyyələri üzrə müraciət sayları

Orta təhsilli şəxslər daha çox müraciət etmişdir.

Təhsil səviyyələri üzrə müraciət sayları	
Müraciət edən vətəndaşın təhsil səviyyəsi	Müraciət edən vətəndaşların sayı
Ümumi təhsil-tam orta təhsil	6744
Ali təhsil-bakalavriat	4041
Ümumi təhsil-ümumi orta təhsil	1888
Tam orta təhsil bazasında orta ixtisas təhsili	1560
Orta ixtisas təhsili	1556
Ümumi orta təhsil bazasında orta ixtisas təhsili	869
İlk nəqə-ixtisas təhsili	777

6. Ən az müraciət edilən 3 filialı tapın



Əlavə

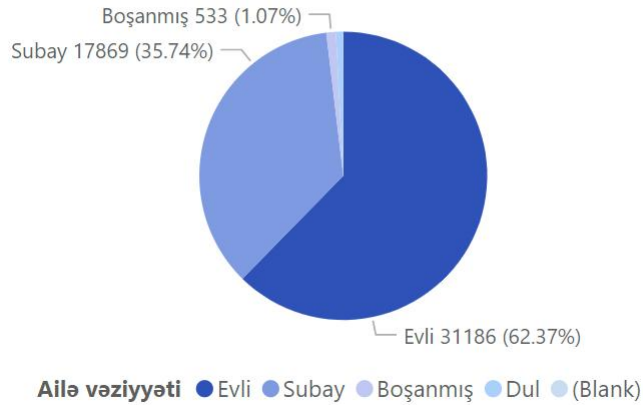
1. Müraciət edilən xidmətlərin sayına baxa bilərik.

Xidmətlərə müraciət edən vətəndaşların sayı	
Müraciət edən vətəndaşa göstərilən xidmət	Müraciət edən vətəndaşların sayı
Haqqı ödənilən ictimai işlər	3964
Özünüməşğulluq (kənd təsərrüfatı zərfləri üzrə)	3936
Peşə yönümlü məsləhət	3905
Əməkhaqqının birgə maliyyələşdirilməsi	3893
Təcrübə proqramlarının təşkili	3888
Əmək yarmarkalarının təşkili	3887
Kvotadan əlavə işadlızalma	3835

2. Müraciət edənlərin ailə vəziyyətləri

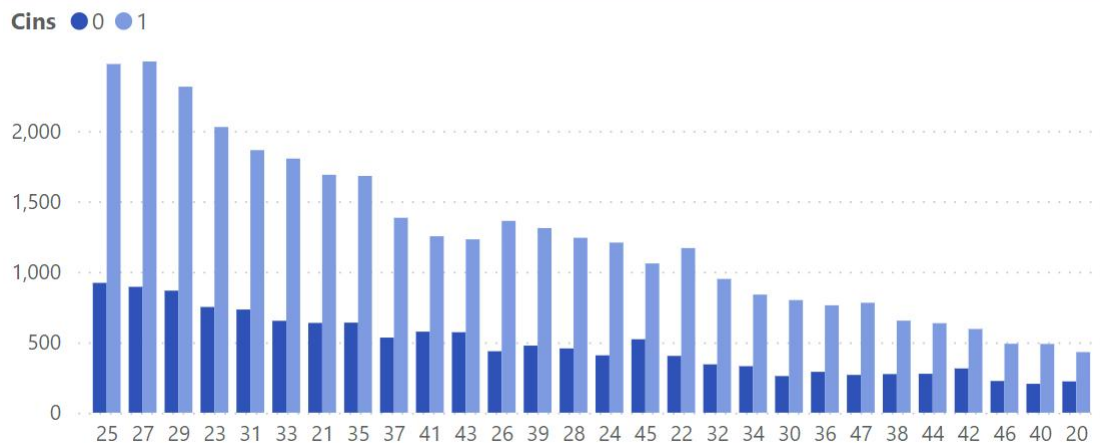
Müraciət edən şəxslərin 62%-i evlidir və 36% -i subaydır. 8 boş sətir var blank onu göstərir.

Müraciət edən vətəndaşların ailə vəziyyəti



3. Qadın və kişilərin hansı yaşda daha çox müraciət etməsi ilə bağlı qrafikə baxaq. Bu qrafikdə kişilərin hər zaman məşğulluğa daha çox müraciət etdiyi görünür. 25 və 27 yaşlı qadın və kişilər daha azdır.

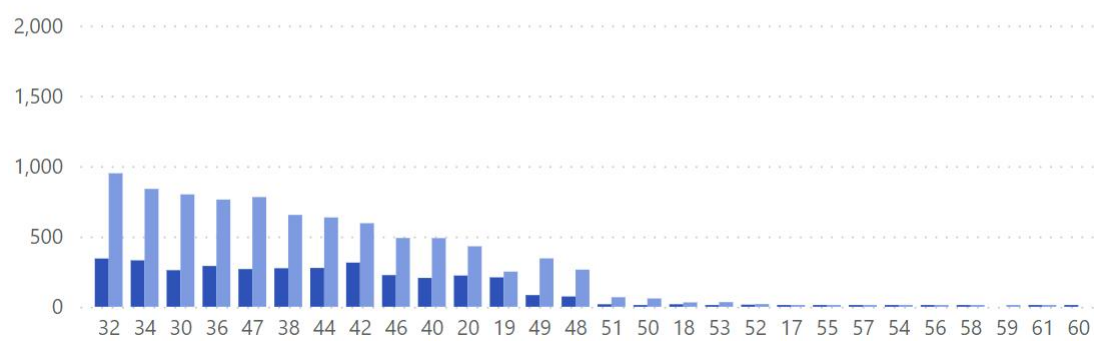
Qadın və kişilərin yaşlara görə müraciət etməsi (0–qadın 1–kişi)



Ən az müraciət edən yaş 60 yaşdır.

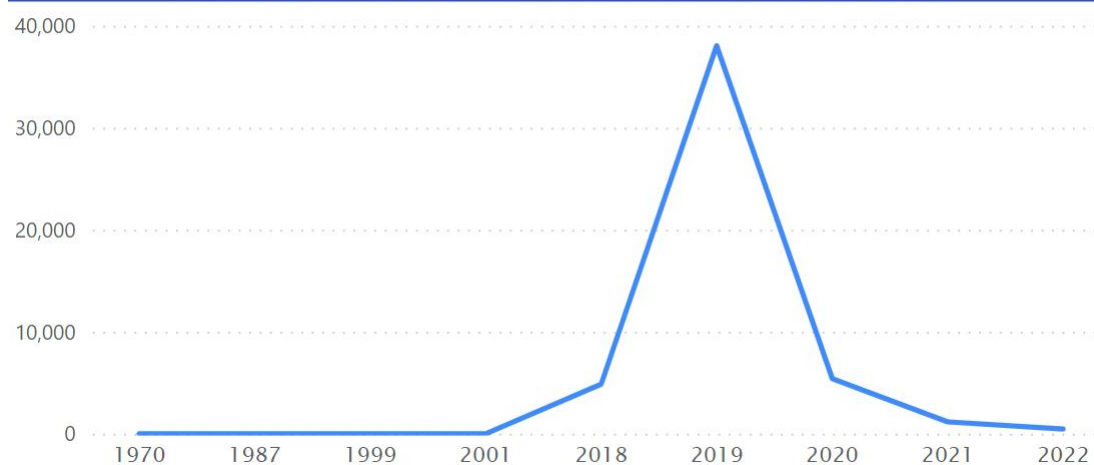
Qadın və kişilərin yaşlara görə müraciət etməsi (0–qadın 1–kişi)

Cins ● 0 ● 1



4. illərboyunca müraciətlərə baxıldığında 2019-cu ildə müraciət sayı kəskin şəkildə artmışdır.

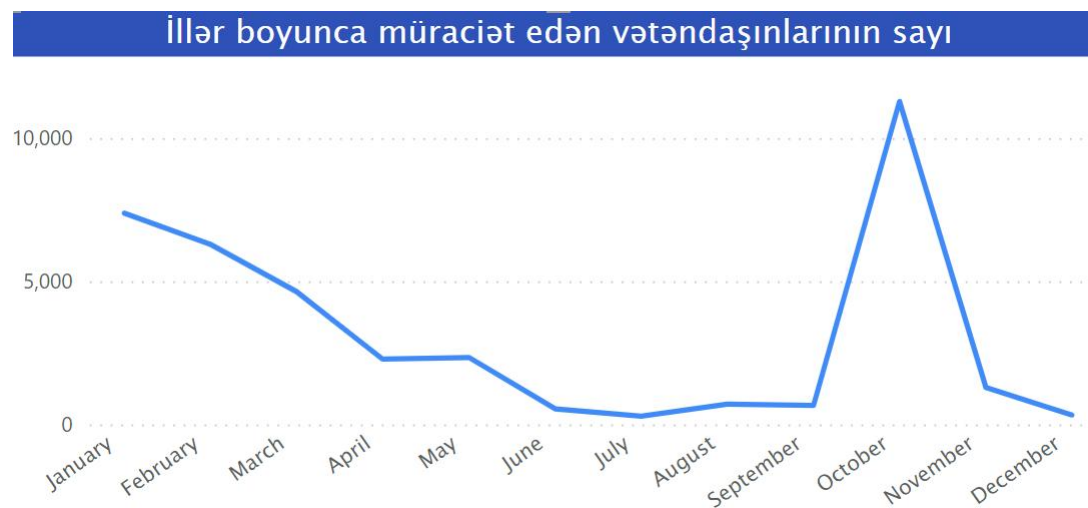
İllər boyunca müraciət edən vətəndaşlarının sayı



2019-cu ildəki müraciətlər göstərirkİ ilin birinci yarısında say çox olmuşdur lakin 3-cü yarısına qədər bu say kəskin şəkildə enmişdir. 3-cü yarısından 4-cü yarısına qədər artmışdır



2019-cu ildə oktyabr və noyabr aylarında müraciət sayı kəskin artmışdır.



Oktyabr ayında olan olan artışı ayın hansı günlərində daha çox olduğuna baxa bilərik.

