

Notebook 1

February 15, 2023

1 MASTER

1.0.1 Matteo Grazioso 884055

1.1 Categorie

```
[1]: # Import libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

```
[2]: # Display all columns and all rows
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)
```

```
[3]: # The data is in the same directory as the notebook and is named 'validazioni.
      ↳txt'.
      # The file contains the data of the validation of tickets in the city of public
      ↳transport of Venice.

      # Import the data into a dataframe of a txt file
      df = pd.read_csv('validazioni.txt', header=0, sep='\t')
```

```
[4]: # Check the first 5 rows of the data
df.head()
```

```
[4]: DATA_VALIDAZIONE      SERIALE  FERMATA  DESCRIZIONE  TITOLO  \
0  13/05/2022 00:00  65676291870913797    5089  FERROVIA "D"    11149
1  13/05/2022 00:00  36141384536591364    5032  FERROVIA "B"    11107
2  13/05/2022 00:00  36144856606063108    5031  P.le Roma "G"    11108
3  13/05/2022 00:00  36144856608393988     506    VENEZIA    12106
4  13/05/2022 00:00  36144856608393732     506    VENEZIA    12106

      DESCRIZIONE_TITOLO
0  7gg-Tp1 43,60-ComVe16,40
```

```

1  48h-Tpl 24,90-ComVe5,10
2  72h-Tpl 33,40-ComVe6,60
3      Bigl Aer-Venezia TSC
4      Bigl Aer-Venezia TSC

```

```

[5]: # Add a new column with the code profile of the ticket
df.insert(6, "TICKET_CODE", 'TBD')

# This column will be filled with the code of the ticket profile according to
# the ticket type and the ticket validity as follows:
# 1. One-day ticket
# 2. Two-day ticket
# 3. Three-day ticket
# 4. Weekly ticket (Seven-day ticket)
# 5. Monthly ticket
# 6. Annual ticket
# 7. Other ticket (if it is necessary to add other types of tickets)

```

```

[29]: # Create a dictionary with the ticket code and the ticket profile
dict_tickets = {'1': 'One-day ticket', '2': 'Two-day ticket', '3': 'Three-day_
               ↪ticket',
               '4': 'Seven-day ticket', '5': 'Monthly ticket', '6': 'Annual_
               ↪ticket',
               '7': 'Other ticket', '7a': 'A_Other ticket', '7b': 'B_Other_
               ↪ticket'}

```

```

[7]: # Which are the unique values of the column 'DESCRIZIONE_TITOLO'?
df['DESCRIZIONE_TITOLO'].unique()

```

```

[7]: array(['7gg-Tpl 43,60-ComVe16,40', '48h-Tpl 24,90-ComVe5,10',
            '72h-Tpl 33,40-ComVe6,60', 'Bigl Aer-Venezia TSC',
            'DAILY PASS VENEZIA - AVM', '72H R.Venice+aerop.AR online',
            '75'-Tpl 6,64-ComVe0,86', 'Biglietto 72 ore Roll. Venice',
            'Bigl.Aut.75'Mestre/Lido-tsc", 'Aer+boat-Tpl14,50-C.Ve1,50',
            '72hAerCS-Tpl40,40-CVe6,60', 'Aeroporto-Venezia AR',
            '72 ore R.Venice online', '7 days online no aerobus',
            '72hAerAR-Tpl46,40-CVe6,60', '48hAerCS-Tpl31,90-CVe5,10',
            'Bicicletta "Palmare"', 'L.17-auto "C"da 4,01 a 4,50 mt',
            '72ore online no aerobus', 'L.17-auto "AeB" fino a 4 metri',
            '48ore online no aerobus', '75'-Tpl 13,28-ComVe1,72",
            'Extra tratta 4 ', '48ore online aerobus AR',
            'Daily Pass Venezia Online', 'Ferry17-carri+35q.rim.',
            'Ferry17-autocarri+35q.', 'Ferry11-autocarri+35q.',
            'Extra tratta 5 ', 'L.17-auto "D" oltre metri 4,50',
            'Extra tratta 2 ', 'Ciclomotore fino 50cc',
            'Bicicletta "biglietteria"', 'Extra tratta 3 ',
            'Bigl.Mestre/Lido 75' a bordo", 'Prenotaz OCCASIONALE si barra',

```

'AtvoCanova+Actv 72Hroll.online', '72 ore R.Venice+aeroporto CS',
 'Extra tratta 1', '7 days online aerobus AR',
 'Atvo Canova+Navig AR online', 'Atvo Canova+Actv 72H online',
 'Tragh-Tpl 4,41-C.Ve0,59', 'ARRIVA VENETO AEROPORTO',
 'ARRIVA VENETO tratta 8-9-10', '24hAerCS-Tpl22,90-CVe5,10',
 'Extra tratta 7 ', 'biglietto merci C.Semplice',
 'Bordo 75min CartaVenezia', 'Bigl.urbano CHIOGGIA',
 '72 ore R.Venice+aeroporto AR', '48hAerAR-Tpl37,90-CVe5,10',
 'Extra tratta 2 TVM', 'Extra tratta 1 TVM',
 'Carnet CHIOGGIA 10c. TICKET', '7ggAerAR-Tpl56,60-CVe16,40',
 "PeopleMover+Bus+Tram 75'", 'Ferry17-Trasporti pericolosi',
 'NA-24H metropolitano ORD+1', '72ore online aerobus CS',
 'ARRIVA VENETO tratta 3', 'abbonamento 30 gg.PeopleMover',
 "ord. navigazione 75' online", 'L.11-auto "D" oltre metri 4,50',
 'Extra tratta 6 ', '24ore online aerobus AR',
 'Cav -Trep + Actv 24H', '7ggAerCS-Tpl50,60-CVe16,40',
 'Ciclomotore oltre 50cc', 'Bagaglio CartaVenezia',
 'Cav - Trep + Actv 72H', 'Cav-Trep - S.Marco AR',
 'Mens. cose animali RETE UNICA', 'Tragh-Tpl 8,82-C.Ve1,18',
 '72H RVenice+aerop.CS online', 'Aeroporto-Venezia CS ONLINE',
 'NA-Traghetto ordinario', 'ARRIVA VENETO tratta 4',
 'Extra tratta 4 TVM', 'ARRIVA VENETO tratta 1',
 '72ore online aerobus AR', '48ore online aerobus CS',
 'NA-24h-Tpl 14,90-Com.Ve5,10', 'Jesolo - S.Marco AR',
 'Jesolo + Actv 24H', "75'-Tpl 12,60-CVe2,40 online",
 'Extra tratta 5 TVM', 'Bicicletta "concessionari"',
 'Gruppi e Scuole', '7 days online aerobus CS',
 'Extra tratta 3 TVM', 'T.Fusina Ve+ACTV 24 ore',
 'T.Fusina Ve+ACTV 72 ore', 'NA-24H metropolitano ORD+2',
 'Tariffa carozzina', 'Aeroporto-Venezia AR ONLINE',
 'ARRIVA Extra tr.8-9-10 BORDO', '24hAerAR-Tpl28,90-CVe5,10',
 'L.11-auto "C"da 4,01 a 4,50 mt', '24H metropolitano ORD online',
 'ARRIVA VENETO tratta 2', 'Extra tratte 2-3-4 BORDO',
 'Bus+People mover online', '24ore online no aerobus',
 'ARRIVAExtra tr.2-3-4 BORDO', 'ARRIVA VENETO tratta 7',
 '24ore online aerobus CS', 'Ev12-Tpl 57,00-C.Ve3,00',
 'ARRIVA Aeroporto BORDO', 'Extra tratta 8-9-10 ',
 'Gruppi e scuole online TVM ar', 'Extra tratta 7 TVM',
 'Extra tratte 5-6-7 BORDO', 'AtvoCanova+Navig 1corsa online',
 'L.11-auto "AeB" fino a 4 metri', 'Gruppi e Scuole AR-SM',
 'Biglietto MOTO FINO 50 cc', 'ARRIVA Extra tr. 5-6-7 BORDO',
 'Ferry11-carri+35q.rim.', "PeopleMover+Bus+Tram 75'carnet",
 'Extra tratta 1 BORDO', 'Atvo Canova+Navig AR', 'Ferry17-AUTOBUS',
 'Atvo Canova+Actv 72H', 'ARRIVA Extra tr.1 BORDO',
 'NA-24H metropolitano ORD.', 'ARRIVA VENETO tratta 6',
 'ARRIVA VENETO tratta 5', '24H metropolitano ORD+1 online',
 'Biglietto Soc. Sportive', '24H metropolitano ORD+2 online',

```
'Ev5-Tpl 33,50-C.Ve1,50', 'Gruppi Organizzati CS',
'Su e Zo per i ponti 2022', 'ARRIVA Integ.Aerop. BORDO',
'Gruppi e scuole online 2viaggi', 'Extra tratta 8-9-10 TVM',
'Traghetto Gratuito', 'Atvo Canova+Navig 1 corsa',
'Gruppi e scuole online TVM cs', 'Ev8-Tpl 45,00-C.Ve3,00',
'Extra tratta 6 TVM', 'NA-24H metropolitano RES+1',
'Ferry11-AUTOBUS', 'Salone Nautico 2022 A/R',
'Salone Nautico 2022 1 corsa', "VENDITA A BORDO 75' CV",
'SpiaggeAR-Tpl 11,75-ComVe1,25', 'Prenotazione Veicolo ABBONATO',
'NA-24H metropolitano RES.', 'Bibione-S.Marco AR',
'Caorle-P.S.Margh-S.Marco AR', 'Caorle-P.S.Margh. + Actv 24H',
'EracleaMare-S.Marco AR', 'Bibione + Actv 24H',
'Lignano-S.Marco AR', 'EracleaMare + Actv 24H',
'NA-24H metropolitano RES+2', 'Extra tratte 8-9-10 BORDO',
'Bigl.urbano CHIOGGIA bordo', 'NA-12h-Tpl 13,40-ComVe4,60',
'Ville Venete+24H actv urb+nav', 'Camp.Marina-S.Marco AR',
"VENDITA A BORDO 75' ORD.", 'Ev3-Tpl 26,50-C.Ve1,50',
'Atvo Canova+Actv 72H rolling', 'Lignano + Actv 24H',
'SOSTITUTIVO Pass Imob'], dtype=object)
```

```
[8]: # Get the number of unique values of the column 'DESCRIZIONE_TITOLO'
num_unique_DESCRIZIONE_TITOLO = len(df['DESCRIZIONE_TITOLO'].unique())
print('The number of unique values of the column DESCRIZIONE_TITOLO is: ',
      num_unique_DESCRIZIONE_TITOLO)
```

The number of unique values of the column DESCRIZIONE_TITOLO is: 168

```
[9]: # How many tickets are there for each type?
df['DESCRIZIONE_TITOLO'].value_counts()
```

```
[9]: DAILY PASS VENEZIA - AVM          972478
      75'-Tpl 6,64-ComVe0,86          743608
      48h-Tpl 24,90-ComVe5,10         600320
      72h-Tpl 33,40-ComVe6,60         492911
      Bigl.Aut.75'Mestre/Lido-tsc     422668
      7gg-Tpl 43,60-ComVe16,40        342870
      75'-Tpl 13,28-ComVe1,72         232644
      Biglietto 72 ore Roll. Venice   170675
      72ore online no aerobus         108357
      7 days online no aerobus        101869
      PeopleMover+Bus+Tram 75'        85968
      Bigl Aer-Venezia TSC            71552
      48ore online no aerobus         65389
      72 ore R.Venice online          59982
      Tragh-Tpl 8,82-C.Ve1,18         38351
      Daily Pass Venezia Online       36709
      Extra tratta 2                  33102
```

Cav -Trep + Actv 24H	25071
Extra tratta 3	22710
Bicicletta "biglietteria"	20041
Jesolo + Actv 24H	17869
NA-24H metropolitano ORD+2	17203
NA-24H metropolitano ORD+1	15662
L.17-auto "D" oltre metri 4,50	14081
Cav-Trep - S.Marco AR	14034
7ggAerAR-Tpl56,60-CVe16,40	13902
Extra tratta 4	13678
Jesolo - S.Marco AR	13620
7 days online aerobus AR	13366
Bigl.Mestre/Lido 75' a bordo	12055
72 ore R.Venice+aeroporto CS	12018
Tragh-Tpl 4,41-C.Ve0,59	11599
Prenotaz OCCASIONALE si barra	11187
48hAerCS-Tpl31,90-CVe5,10	9889
72hAerCS-Tpl40,40-CVe6,60	9855
L.17-auto "AeB" fino a 4 metri	8279
72hAerAR-Tpl46,40-CVe6,60	8152
24hAerCS-Tpl22,90-CVe5,10	8048
ord. navigazione 75' online	6976
L.17-auto "C"da 4,01 a 4,50 mt	6941
Extra tratta 1	6800
Aer+boat-Tpl14,50-C.Ve1,50	6693
Aeroporto-Venezia AR	6520
ARRIVA VENETO tratta 8-9-10	6323
7ggAerCS-Tpl50,60-CVe16,40	6295
72H R.Venice+aerop.CS online	6190
72ore online aerobus AR	6055
Extra tratta 5	5347
72H R.Venice+aerop.AR online	5087
72 ore R.Venice+aeroporto AR	4963
72ore online aerobus CS	4565
Bicicletta "Palmare"	4307
Carnet CHIOGGIA 10c. TICKET	4241
7 days online aerobus CS	3933
T.Fusina Ve+ACTV 24 ore	3713
75'-Tpl 12,60-CVe2,40 online	3524
Bigl.urbano CHIOGGIA	3430
Gruppi e Scuole	3421
Bordo 75min CartaVenezia	3345
48hAerAR-Tpl37,90-CVe5,10	3312
Gruppi e scuole online TVM ar	3097
Ferry17-autocarri+35q.	3037
Salone Nautico 2022 A/R	2916
Atvo Canova+Actv 72H online	2788

T.Fusina Ve+ACTV 72 ore	2691
Tariffa carrozzina	2573
48ore online aerobus CS	2386
AtvoCanova+Actv 72Hroll.online	2260
Bicicletta "concessionari"	2141
Ciclomotore fino 50cc	2101
Su e Zo per i ponti 2022	1998
Gruppi e Scuole AR-SM	1907
L.11-auto "D" oltre metri 4,50	1845
Extra tratta 6	1815
Atvo Canova+Actv 72H	1769
Aeroporto-Venezia AR ONLINE	1697
48ore online aerobus AR	1532
abbonamento 30 gg.PeopleMover	1509
Caorle-P.S.Margh-S.Marco AR	1484
L.11-auto "AeB" fino a 4 metri	1450
Bus+People mover online	1339
Caorle-P.S.Margh. + Actv 24H	1297
24hAerAR-Tpl28,90-CVe5,10	1209
Bibione-S.Marco AR	1173
L.11-auto "C"da 4,01 a 4,50 mt	1109
ARRIVA VENETO tratta 1	1090
Ev3-Tpl 26,50-C.Ve1,50	1000
24ore online aerobus CS	893
Bibione + Actv 24H	845
Ferry11-autocarri+35q.	838
Extra tratta 1 TVM	801
Extra tratta 8-9-10	766
Aeroporto-Venezia CS ONLINE	749
Extra tratta 2 TVM	748
Salone Nautico 2022 1 corsa	718
Ferry17-carri+35q.rim.	668
PeopleMover+Bus+Tram 75'carnet	648
Extra tratte 2-3-4 BORDO	641
EracleaMare-S.Marco AR	606
ARRIVA VENETO tratta 4	580
Ev5-Tpl 33,50-C.Ve1,50	558
Cav - Trep + Actv 72H	544
NA-24h-Tpl 14,90-Com.Ve5,10	532
Extra tratta 7	514
ARRIVA VENETO tratta 6	509
24ore online no aerobus	457
Ev12-Tpl 57,00-C.Ve3,00	455
24H metropolitano ORD online	435
Extra tratta 3 TVM	398
Extra tratta 4 TVM	387
24H metropolitano ORD+2 online	346

Atvo Canova+Navig 1 corsa	345
Ev8-Tpl 45,00-C.Ve3,00	308
Gruppi Organizzati CS	303
ARRIVA VENETO AEROPORTO	298
ARRIVA VENETO tratta 7	279
Biglietto Soc. Sportive	274
Lignano-S.Marco AR	274
Bagaglio Carta Venezia	270
Gruppi e scuole online TVM cs	264
Atvo Canova+Navig AR	257
Biglietto MOTO FINO 50 cc	244
24H metropolitano ORD+1 online	239
Ciclomotore oltre 50cc	238
ARRIVA VENETO tratta 3	237
ARRIVA Extra tr.8-9-10 BORDO	236
24ore online aerobus AR	222
ARRIVA VENETO tratta 2	206
Extra tratta 1 BORDO	205
Eraclea Mare + Actv 24H	204
ARRIVA Extra tr.1 BORDO	182
Gruppi e scuole online 2viaggi	180
ARRIVA Extra tr.2-3-4 BORDO	171
ARRIVA Extra tr. 5-6-7 BORDO	166
Extra tratta 5 TVM	160
biglietto merci C.Semplice	159
Ferry17-AUTOBUS	134
SpiaggeAR-Tpl 11,75-ComVe1,25	125
NA-Traghetto ordinario	109
Atvo Canova+Navig AR online	102
NA-24H metropolitano ORD.	97
AtvoCanova+Navig 1corsa online	80
ARRIVA VENETO tratta 5	74
Mens. cose animali RETE UNICA	48
Extra tratte 5-6-7 BORDO	38
Ferry17-Trasporti pericolosi	32
NA-24H metropolitano RES+1	31
Lignano + Actv 24H	25
Extra tratta 8-9-10 TVM	23
ARRIVA Aeroporto BORDO	22
Prenotazione Veicolo ABBONATO	20
Ville Venete+24H actv urb+nav	18
Extra tratta 7 TVM	18
NA-24H metropolitano RES+2	17
ARRIVA Integ.Aerop. BORDO	14
NA-24H metropolitano RES.	13
VENDITA A BORDO 75' CV	13
Ferry11-carri+35q.rim.	13

Bigl.urbano CHIoggia bordo	12
VENDITA A BORDO 75' ORD.	10
Extra tratta 6 TVM	9
Traghetto Gratuito	8
Ferry11-AUTOBUS	3
Extra tratte 8-9-10 BORDO	3
NA-12h-Tpl 13,40-ComVe4,60	3
Camp.Marina-S.Marco AR	2
Atvo Canova+Actv 72H rolling	1
SOSTITUTIVO Pass Imob	1
Name: DESCRIZIONE_TITOLO, dtype: int64	

1.1.1 One-day tickets

```
[10]: # Which type of ticket are one-day tickets and how many are there?
# Exclude the tickets that contains also 48h or 72h
df[df['DESCRIZIONE_TITOLO'].str.contains('giornaliero|24h|24_
ore|24|Daily|daily|DAILY') & ~df['DESCRIZIONE_TITOLO'].str.
contains('48|72')]['DESCRIZIONE_TITOLO'].value_counts()
```

DAILY PASS VENEZIA - AVm	972478
Daily Pass Venezia Online	36709
Cav -Trep + Actv 24H	25071
Jesolo + Actv 24H	17869
NA-24H metropolitano ORD+2	17203
NA-24H metropolitano ORD+1	15662
24hAerCS-Tpl22,90-CVe5,10	8048
T.Fusina Ve+ACTV 24 ore	3713
Caorle-P.S.Margh. + Actv 24H	1297
24hAerAR-Tpl28,90-CVe5,10	1209
24ore online aerobus CS	893
Bibione + Actv 24H	845
NA-24h-Tpl 14,90-Com.Ve5,10	532
24ore online no aerobus	457
24H metropolitano ORD online	435
24H metropolitano ORD+2 online	346
24H metropolitano ORD+1 online	239
24ore online aerobus AR	222
EracleaMare + Actv 24H	204
NA-24H metropolitano ORD.	97
NA-24H metropolitano RES+1	31
Lignano + Actv 24H	25
Ville Venete+24H actv urb+nav	18
NA-24H metropolitano RES+2	17
NA-24H metropolitano RES.	13
Name: DESCRIZIONE_TITOLO, dtype: int64	


```
[11]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
      ↪ according to the ticket type and the ticket validity
df.loc[df['DESCRIZIONE_TITOLO'].str.contains('giornaliero|24h|24_
      ↪ ore|24|Daily|daily|DAILY') & ~df['DESCRIZIONE_TITOLO'].str.
      ↪ contains('48|72'), 'TICKET_CODE'] = '1'
```

1.1.2 Two days tickets

```
[12]: # Which type of ticket are two-day tickets and how many are there?
df[df['DESCRIZIONE_TITOLO'].str.contains('48')]['DESCRIZIONE_TITOLO'].
    ↪ value_counts()
```

```
[12]: 48h-Tp1 24,90-ComVe5,10      600320
      48ore online no aerobus      65389
      48hAerCS-Tp131,90-CVe5,10    9889
      48hAerAR-Tp137,90-CVe5,10    3312
      48ore online aerobus CS      2386
      48ore online aerobus AR      1532
      Name: DESCRIZIONE_TITOLO, dtype: int64
```

```
[13]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
      ↪ according to the ticket type and the ticket validity
df.loc[df['DESCRIZIONE_TITOLO'].str.contains('48'), 'TICKET_CODE'] = '2'
```

1.1.3 Three days tickets

```
[14]: # Which type of ticket are three-day tickets and how many are there?
df[df['DESCRIZIONE_TITOLO'].str.contains('72')]['DESCRIZIONE_TITOLO'].
    ↪ value_counts()
```

```
[14]: 72h-Tp1 33,40-ComVe6,60      492911
      75'-Tp1 13,28-ComVe1,72      232644
      Biglietto 72 ore Roll. Venice 170675
      72ore online no aerobus      108357
      72 ore R.Venice online      59982
      72 ore R.Venice+aeroporto CS 12018
      72hAerCS-Tp140,40-CVe6,60    9855
      72hAerAR-Tp146,40-CVe6,60    8152
      72H R.Venice+aerop.CS online 6190
      72ore online aerobus AR      6055
      72H R.Venice+aerop.AR online 5087
      72 ore R.Venice+aeroporto AR 4963
      72ore online aerobus CS      4565
      Atvo Canova+Actv 72H online  2788
      T.Fusina Ve+ACTV 72 ore      2691
      AtvoCanova+Actv 72Hroll.online 2260
```

```
Atvo Canova+Actv 72H          1769
Cav - Trep + Actv 72H         544
Atvo Canova+Actv 72H rolling   1
Name: DESCRIZIONE_TITOLO, dtype: int64
```

```
[15]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
      ↪ according to the ticket type and the ticket validity
      df.loc[df['DESCRIZIONE_TITOLO'].str.contains('72'), 'TICKET_CODE'] = '3'
```

1.1.4 Seven days tickets

```
[16]: # Which type of ticket are weekly tickets and how many are there?
      # Exclude the tickets that contains also 72, 75 that are three-day tickets, 17,
      ↪ 48h, 57 that are other types of tickets and
      # 'tratt*' and 'tr' that are reserved to specific routes
      df[df['DESCRIZIONE_TITOLO'].str.contains('settimanale|7|Weekly')] &
      ↪ ~df['DESCRIZIONE_TITOLO'].str.
      ↪ contains('72|75|17|48h|57|tratt*|tr')]['DESCRIZIONE_TITOLO'].value_counts()
```

```
[16]: 7gg-Tp1 43,60-ComVe16,40      342870
      7 days online no aerobus      101869
      7ggAerAR-Tp156,60-CVe16,40    13902
      7 days online aerobus AR      13366
      7ggAerCS-Tp150,60-CVe16,40    6295
      7 days online aerobus CS      3933
      Name: DESCRIZIONE_TITOLO, dtype: int64
```

```
[17]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
      ↪ according to the ticket type and the ticket validity
      df.loc[df['DESCRIZIONE_TITOLO'].str.contains('settimanale|7|Weekly')] &
      ↪ ~df['DESCRIZIONE_TITOLO'].str.contains('72|75|17|48h|57|tratt*|tr'),
      ↪ 'TICKET_CODE'] = '4'
```

1.1.5 Monthly tickets

```
[18]: # Which type of ticket are monthly tickets and how many are there?
      df[df['DESCRIZIONE_TITOLO'].str.
      ↪ contains('abbonamento|mensile|30')]['DESCRIZIONE_TITOLO'].value_counts()
```

```
[18]: abbonamento 30 gg.PeopleMover    1509
      Name: DESCRIZIONE_TITOLO, dtype: int64
```

```
[19]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
      ↪ according to the ticket type and the ticket validity
      df.loc[df['DESCRIZIONE_TITOLO'].str.contains('abbonamento|mensile|30'),
      ↪ 'TICKET_CODE'] = '5'
```

1.1.6 Yearly tickets

```
[20]: # Which type of ticket are annual tickets and how many are there?
df[df['DESCRIZIONE_TITOLO'].str.
    ↪contains('annuale|365|year|yearly')]['DESCRIZIONE_TITOLO'].value_counts()
```

```
[20]: Series([], Name: DESCRIZIONE_TITOLO, dtype: int64)
```

```
[21]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
    ↪according to the ticket type and the ticket validity
df.loc[df['DESCRIZIONE_TITOLO'].str.contains('annuale|365|year|yearly'),
    ↪'TICKET_CODE'] = '6'
```

1.1.7 75 minutes tickets

```
[22]: # Which type of ticket are 75' (75 minutes) tickets and how many are there?
df[df['DESCRIZIONE_TITOLO'].str.contains('75')]['DESCRIZIONE_TITOLO'].
    ↪value_counts()
```

```
[22]: 75'-Tpl 6,64-ComVe0,86          743608
Bigl.Aut.75'Mestre/Lido-tsc      422668
75'-Tpl 13,28-ComVe1,72          232644
PeopleMover+Bus+Tram 75'        85968
Bigl.Mestre/Lido 75' a bordo     12055
ord. navigazione 75' online      6976
75'-Tpl 12,60-CVe2,40 online     3524
Bordo 75min CartaVenezia        3345
PeopleMover+Bus+Tram 75'carnet   648
SpiaggeAR-Tpl 11,75-ComVe1,25    125
VENDITA A BORDO 75' CV           13
VENDITA A BORDO 75' ORD.         10
Name: DESCRIZIONE_TITOLO, dtype: int64
```

```
[23]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
    ↪according to the ticket type and the ticket validity
df.loc[df['DESCRIZIONE_TITOLO'].str.contains('75'), 'TICKET_CODE'] = '7a'
```

1.1.8 Other types of tickets

```
[24]: # Which type of ticket are other tickets and how many are there?
# The other tickets are the tickets that are not already classified in the
    ↪previous categories
df[~df['TICKET_CODE'].isin(['1','2','3','4','5','6',
    ↪'7a'])]['DESCRIZIONE_TITOLO'].value_counts()
```

[24]: Bigl Aer-Venezia TSC	71552
Tragh-Tpl 8,82-C.Ve1,18	38351
Extra tratta 2	33102
Extra tratta 3	22710
Bicicletta "biglietteria"	20041
L.17-auto "D" oltre metri 4,50	14081
Cav-Trep - S.Marco AR	14034
Extra tratta 4	13678
Jesolo - S.Marco AR	13620
Tragh-Tpl 4,41-C.Ve0,59	11599
Prenotaz OCCASIONALE si barra	11187
L.17-auto "AeB" fino a 4 metri	8279
L.17-auto "C"da 4,01 a 4,50 mt	6941
Extra tratta 1	6800
Aer+boat-Tpl14,50-C.Ve1,50	6693
Aeroporto-Venezia AR	6520
ARRIVA VENETO tratta 8-9-10	6323
Extra tratta 5	5347
Bicicletta "Palmare"	4307
Carnet CHIOGGIA 10c. TICKET	4241
Bigl.urbano CHIOGGIA	3430
Gruppi e Scuole	3421
Gruppi e scuole online TVM ar	3097
Ferry17-autocarri+35q.	3037
Salone Nautico 2022 A/R	2916
Tariffa carrozzina	2573
Bicicletta "concessionari"	2141
Ciclomotore fino 50cc	2101
Su e Zo per i ponti 2022	1998
Gruppi e Scuole AR-SM	1907
L.11-auto "D" oltre metri 4,50	1845
Extra tratta 6	1815
Aeroporto-Venezia AR ONLINE	1697
Caorle-P.S.Margh-S.Marco AR	1484
L.11-auto "AeB" fino a 4 metri	1450
Bus+People mover online	1339
Bibione-S.Marco AR	1173
L.11-auto "C"da 4,01 a 4,50 mt	1109
ARRIVA VENETO tratta 1	1090
Ev3-Tpl 26,50-C.Ve1,50	1000
Ferry11-autocarri+35q.	838
Extra tratta 1 TVM	801
Extra tratta 8-9-10	766
Aeroporto-Venezia CS ONLINE	749
Extra tratta 2 TVM	748
Salone Nautico 2022 1 corsa	718
Ferry17-carri+35q.rim.	668

Extra tratte 2-3-4 BORDO	641
EracleaMare-S.Marco AR	606
ARRIVA VENETO tratta 4	580
Ev5-Tp1 33,50-C.Ve1,50	558
Extra tratta 7	514
ARRIVA VENETO tratta 6	509
Ev12-Tp1 57,00-C.Ve3,00	455
Extra tratta 3 TVM	398
Extra tratta 4 TVM	387
Atvo Canova+Navig 1 corsa	345
Ev8-Tp1 45,00-C.Ve3,00	308
Gruppi Organizzati CS	303
ARRIVA VENETO AEROPORTO	298
ARRIVA VENETO tratta 7	279
Lignano-S.Marco AR	274
Biglietto Soc. Sportive	274
Bagaglio CartaVenezia	270
Gruppi e scuole online TVM cs	264
Atvo Canova+Navig AR	257
Biglietto MOTO FINO 50 cc	244
Ciclomotore oltre 50cc	238
ARRIVA VENETO tratta 3	237
ARRIVA Extra tr.8-9-10 BORDO	236
ARRIVA VENETO tratta 2	206
Extra tratta 1 BORDO	205
ARRIVA Extra tr.1 BORDO	182
Gruppi e scuole online 2viaggi	180
ARRIVAExtra tr.2-3-4 BORDO	171
ARRIVA Extra tr. 5-6-7 BORDO	166
Extra tratta 5 TVM	160
biglietto merci C.Semplice	159
Ferry17-AUTOBUS	134
NA-Traghetto ordinario	109
Atvo Canova+Navig AR online	102
AtvoCanova+Navig 1corsa online	80
ARRIVA VENETO tratta 5	74
Mens. cose animali RETE UNICA	48
Extra tratte 5-6-7 BORDO	38
Ferry17-Trasporti pericolosi	32
Extra tratta 8-9-10 TVM	23
ARRIVA Aeroporto BORDO	22
Prenotazione Veicolo ABBONATO	20
Extra tratta 7 TVM	18
ARRIVA Integ.Aerop. BORDO	14
Ferry11-carri+35q.rim.	13
Bigl.urbano CHIOGGIA bordo	12
Extra tratta 6 TVM	9

Traghetto Gratuito	8
Ferry11-AUTOBUS	3
Extra tratte 8-9-10 BORDO	3
NA-12h-Tp1 13,40-ComVe4,60	3
Camp.Marina-S.Marco AR	2
SOSTITUTIVO Pass Imob	1

Name: DESCRIZIONE_TITOLO, dtype: int64

```
[25]: # Populate the column 'TICKET_CODE' with the code of the ticket profile
      ↪ according to the ticket type and the ticket validity
df.loc[~df['TICKET_CODE'].isin(['1','2','3','4','5','6','7a']), 'TICKET_CODE']
      ↪= '7b'
```

1.1.9 Summary of the ticket profiles

```
[54]: # Print the number of tickets for each ticket profile code ordered by the code
      ↪ of the ticket profile; print the name of the ticket profile using the
      ↪ dictionary 'dict_tickets'
df['TICKET_CODE'].value_counts().sort_index().rename(dict_tickets).
      ↪reindex(dict_tickets.values(), fill_value=0)
```

```
[54]: One-day ticket      1103633
      Two-day ticket      682828
      Three-day ticket     898863
      Seven-day ticket     482235
      Monthly ticket       1509
      Annual ticket         0
      Other ticket          0
      A_Other ticket      1511584
      B_Other ticket      375989
      Name: TICKET_CODE, dtype: int64
```

```
[59]: # Countplot of the column 'TICKET_CODE'
fig, ax = plt.subplots(figsize=(15,8))
# Countplot of the column 'TICKET_CODE'
sns.countplot(x='TICKET_CODE', data=df, order=df['TICKET_CODE'].value_counts().
      ↪sort_index().index)
plt.title('Countplot of the column TICKET_CODE', fontsize=20)
plt.xlabel('Ticket code', fontsize=15)
plt.ylabel('Count (in millions)', fontsize=15)

# Change yticks to have a better visualization
scale = np.arange(0, max(df['TICKET_CODE'].value_counts())+100000, 100000)
plt.yticks(scale)

# Add the percentage of each category on top of the bars
```

```

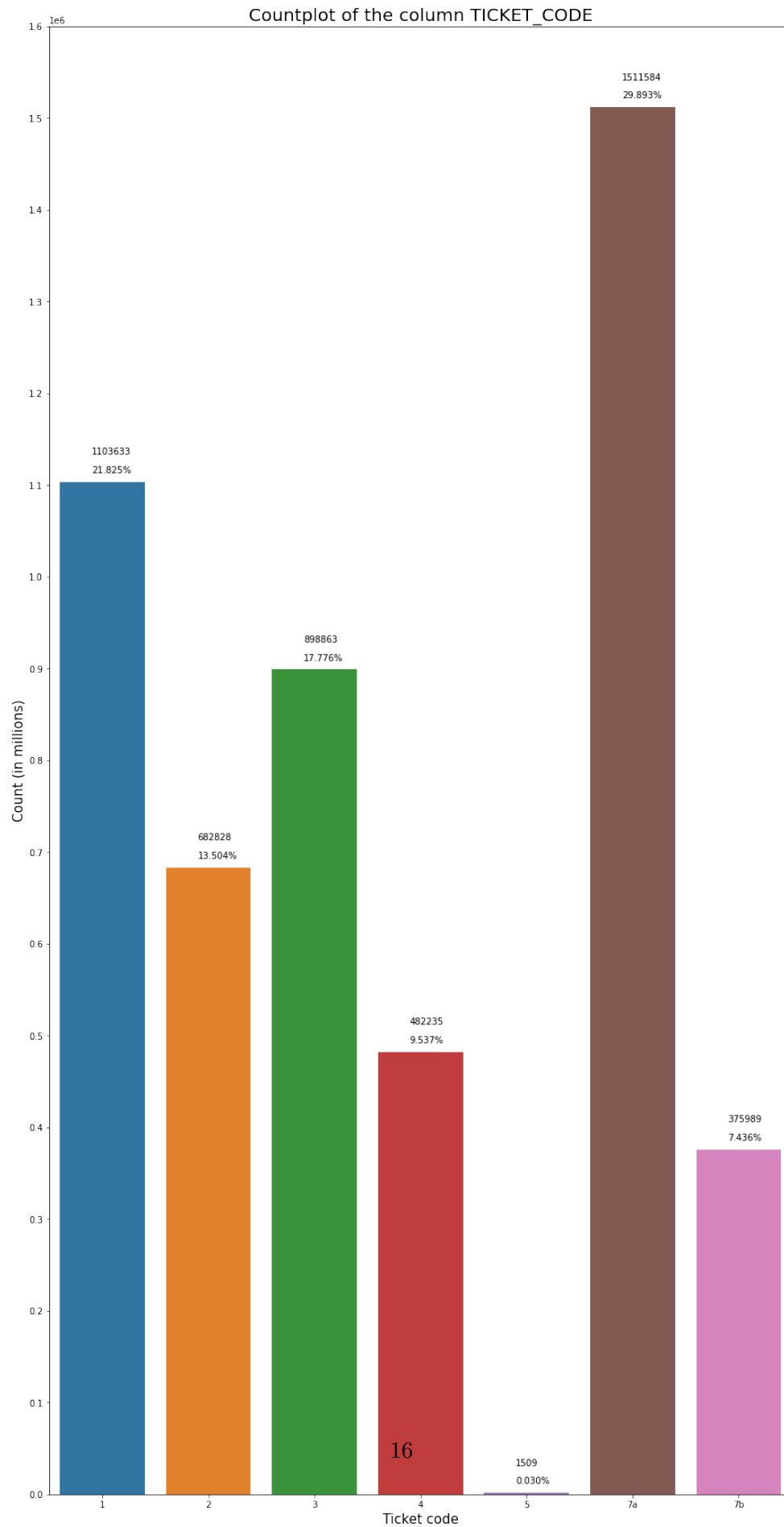
for p in ax.patches:
    ax.annotate('{:.3f}%'.format(100*p.get_height()/len(df)), (p.get_x()+0.3, p.
↳get_height()+10000))

# Add the count of each category on top of the bars
for p in ax.patches:
    ax.annotate('{:.0f}'.format(p.get_height()), (p.get_x()+0.3, p.
↳get_height()+30000))

# Add a padding on the top of the plot
plt.subplots_adjust(top=3)

plt.show()

```




```
[65]: # Plot a pie chart of the column 'TICKET_CODE'
fig, ax = plt.subplots(figsize=(20,10))
df['TICKET_CODE'].value_counts().sort_index().plot.pie(autopct='%1.3f%%',
    ↪startangle=90, pctdistance=0.85, labeldistance=1.1)

# Add the name of the ticket profile on the pie chart
plt.legend(labels=df['TICKET_CODE'].value_counts().sort_index().
    ↪rename(dict_tickets).index, loc='center left', bbox_to_anchor=(1, 0, 0.5,
    ↪1), fontsize=15)

plt.title('Pie chart of the column TICKET_CODE', fontsize=20)
plt.ylabel('')
fig.patch.set_facecolor('white')
plt.show()
```

Pie chart of the column TICKET_CODE

