Για την περάτωση της άσκησης, αρχικά κατέβασα το αρχείο [finance\_liquor\_sales.sql](https://github.com/Workearly/Final-Assignment/blob/main/finance_liquor_sales.sql) και το άνοιξα με το πρόγραμμα MySQL Workbench.  
Στη συνέχεια χρησιμοποίησα αυτό το Query για να εξάγω τις στήλες από τα tables μεταξύ 2016 και 2019:

SELECT date, address, store\_number, store\_name, invoice\_and\_item\_number, zip\_code, bottles\_sold, item\_description

FROM `liquorsales`.`finance\_liquor\_sales`

WHERE date BETWEEN '2016/01/01' and '2019/12/31'

group by zip\_code

having count(\*) >1

or count(\*) =1;

Έπειτα έκανα εξαγωγή το αποτέλεσμα σε csv αρχείο και το εισήγαγα στο πρόγραμμα PyCharm.

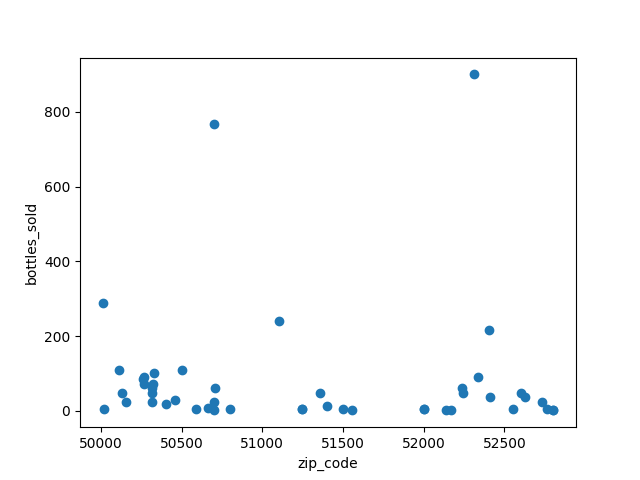
Χρησιμοποιώντας τις παρακάτω εντολές βρήκα τα αποτελέσματα από τα πιο δημοφιλή ποτά προς πώληση για τον κάθε ταχυδρομικό κώδικα και βρήκα το αποτέλεσμα από το μερίδιο της αγοράς που αντιστοιχεί στο κάθε ποτό αντίστοιχα:

import pandas as pd  
import pandas  
import numpy as np  
import matplotlib.pyplot as plt  
  
df = pd.read\_csv('2016-2019.csv')  
pandas.set\_option('display.max\_rows', df.shape[0]+1)  
  
df.drop\_duplicates(subset='zip\_code', inplace=True, keep='last')  
  
cnz = df.groupby(['zip\_code','bottles\_sold'])  
  
print(cnz.first())  
  
df.arr1 = np.array('bottles\_sold')  
print(df.sum())  
  
ab = np.array([288,4,108,48,24,84,72,90,60,48,24,72,102,18,30,108,4,8,2,768,24,60,5,240,5,6,48,12,4,2,4,5,  
 2,1,60,48,900,90,216,36,6,48,36,24,4,3,2])/ 3853\*100

new\_ab = list(np.around(np.array(ab),2))  
print(new\_ab)

x = np.array([50010,50022,50111,50131,50158,50263,50265,50266,50314,50316,50317,50320,50327,50401,50461,50501,  
 50588,50662,50701,50702,50703,50707,50801,51106,51246,51247,51360,51401,51501,51555,52001,52003,52136,  
 52172,52240,52241,52314,52338,52402,52411,52556,52601,52627,52732,52761,52803,52804])  
  
y = np.array([288,4,108,48,24,84,72,90,60,48,24,72,102,18,30,108,4,8,2,768,24,60,5,240,5,6,48,12,4,2,4,5,  
 2,1,60,48,900,90,216,36,6,48,36,24,4,3,2])  
  
plt.xlabel('zip\_code')  
plt.ylabel('bottles\_sold')  
  
plt.scatter(x,y)  
  
plt.show()

Με αυτόν τον τρόπο παρουσιάζεται το σχεδιάγραμμα στο Matoplotlib όπως και στο παράδειγμα.



Μια από τις δυσκολίες που συνάντησα είναι μετά τις εντολές στο πρόγραμμα PyCharm το αποτέλεσμα που βγαίνει κάτω στο παράθυρο του run είναι εφικτό να αποθηκευτεί σε μορφή csv ώστε να το χρησιμοποιήσω στο πρόγραμμα Tableau;

Επίσης, μια επιπρόσθετη παρόμοια δυσκολία που μου παρουσιάστηκε είναι ότι αφού χρησιμοποίησα τις εντολές για να βρω τα πιο δημοφιλή μπουκάλια που πουλήθηκαν στον κάθε ταχυδρομικό κώδικα αντίστοιχα, βρίσκοντας το αποτέλεσμα μετά από μια σειρά εντολών μέσω της Python και Pandas, αδυνατούσα να χρησιμοποιήσω τη λίστα που έβρισκα με αποτέλεσμα να γράψω ένα προς ένα τα δεδομένα του x και του y στο σχεδιάγραμμα της Matplotlib.

Δηλαδή η βασική απορία μου είναι αν μπορώ εξάγω και αν ναι πως τα αποτελέσματα από το PyCharm σε μορφή csv