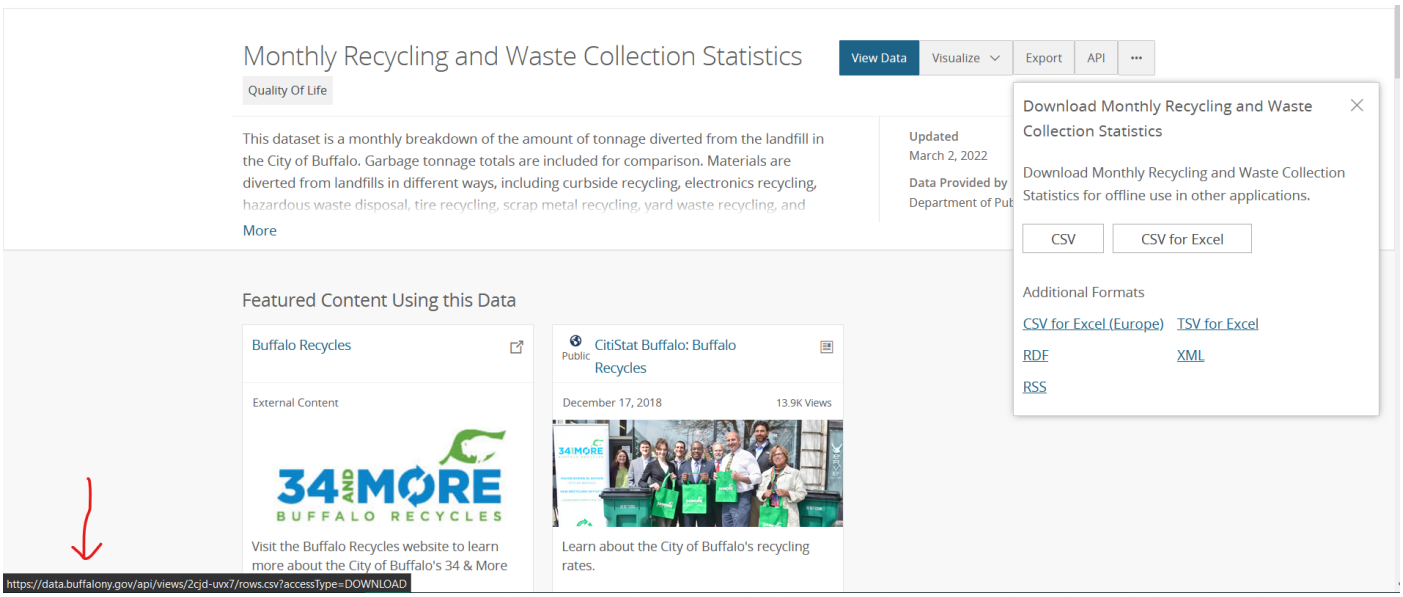
# ΑΡΧΕΣ ΓΛΩΣΣΩΝ

# ΠΡΟΑΙΡΕΤΙΚΗ ΑΣΚΗΣΗ PYTHON

Όνομα: Γιακαλλής Κωνσταντίνος ΑΜ: 1072533 Έτος:3ο

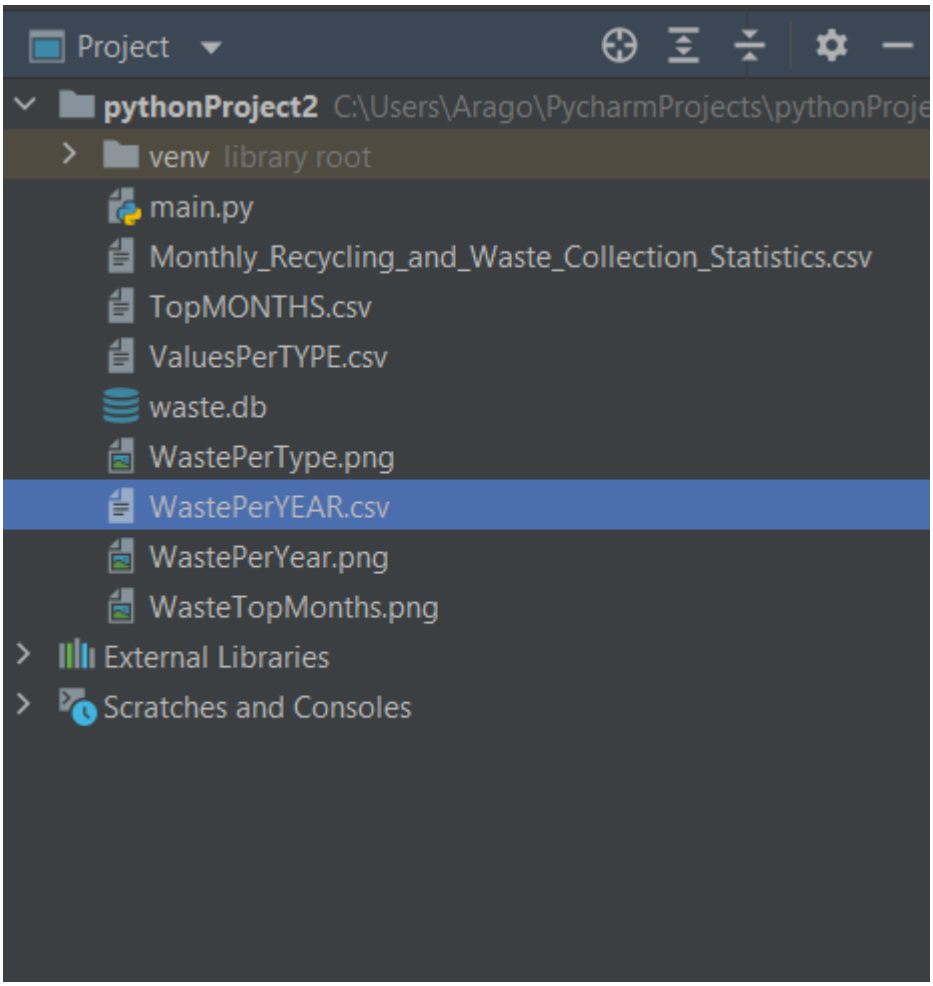
Για την περισυλλογή του csv απο την σελίδα απλά αντέγραψα το link που εμφανίζεται όταν κάνεις hover πάνω απο το download button στην σελίδα.

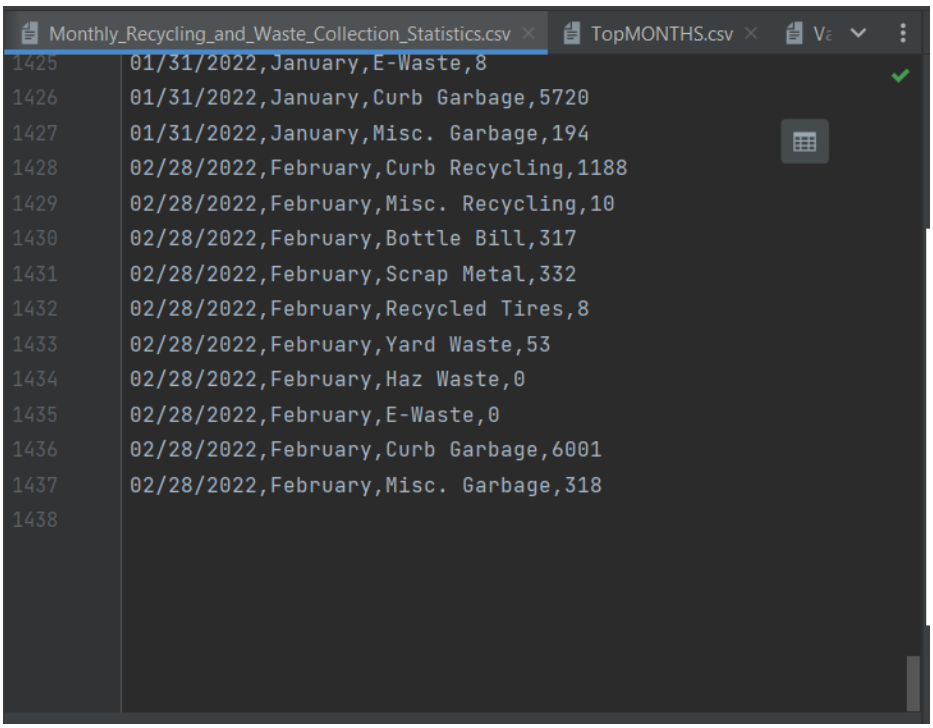



Για να κατεβάσω το csv και να το τρέξω χρησιμοποιήθηκε η βιβλιοθήκη requests. Το πρόγραμμα παίρνει τα contents του download, δημιουργεί ένα csv και τα αντιγράφει μέσα. (βλ. Γραμμές 6-11 στον κώδικα).

Έπειτα φτιάχνουμε ένα dataframe για να μπορέσουμε να επεξεργαστούμε τα δεδομένα του csv και να κάνουμε διάφορα operations όπως group by. Σε αυτό μας βοηθά η βιβλιοθήκη pandas.Επίσης σώζουμε όλα τα νέα δεδομένα σε csv. (βλ. Γραμμές 13-31)

Κάποιες εικόνες των csv:



Statistics.csv | TopMONTHS.csv | ValuesPerTYPE.csv | WastePerYEAR.csv

| | DATE | TOTAL (IN TONS) |
|---|---|---|
| 1 | 2011 | 173428 |
| 2 | 2012 | 203032 |
| 3 | 2013 | 209352 |
| 4 | 2014 | 214605 |
| 5 | 2015 | 156190 |
| 6 | 2016 | 147764 |
| 7 | 2017 | 146914 |
| 8 | 2018 | 149180 |
| 9 | 2019 | 152221 |
| 10 | 2020 | 152815 |
| 11 | 2021 | 153065 |
| 12 | 2022 | 16008 |

Statistics.csv | TopMONTHS.csv | ValuesPerTYPE.csv | W

| | TYPE | TOTAL (IN TONS) |
|---|---|---|
| 1 | Asphalt Debris | 208000 |
| 2 | Bottle Bill | 42358 |
| 3 | Curb Garbage | 1074974 |
| 4 | Curb Recycling | 170338 |
| 5 | E-Waste | 3792 |
| 6 | Haz Waste | 130 |
| 7 | Misc. Garbage | 179894 |
| 8 | Misc. Recycling | 41428 |
| 9 | Recycled Tires | 3222 |
| 10 | Scrap Metal | 38171 |
| 11 | Sidewalk Debris | 18908 |
| 12 | Yard Waste | 93359 |

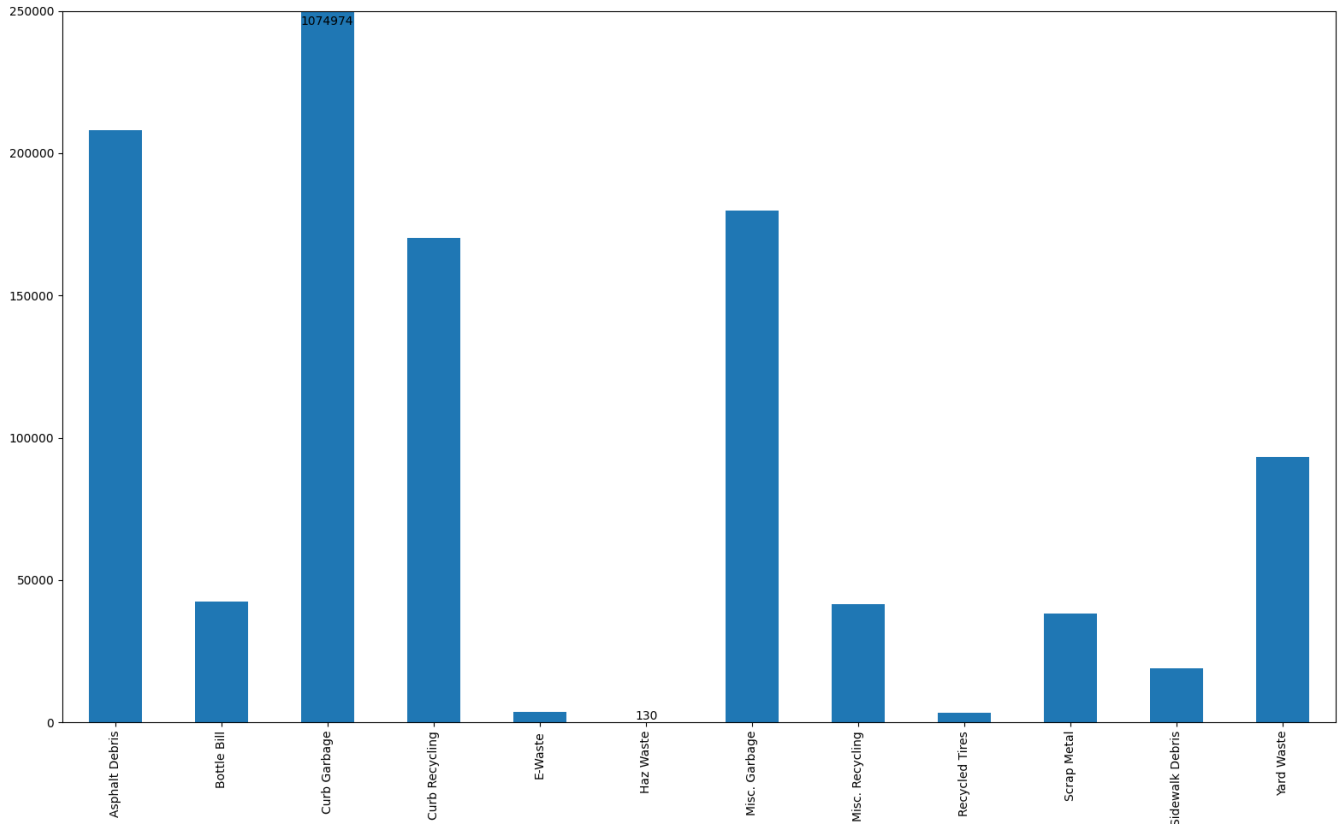
Statistics.csv | TopMONTHS.csv | ValuesPerTYPE.csv | WastePerYEAR.csv

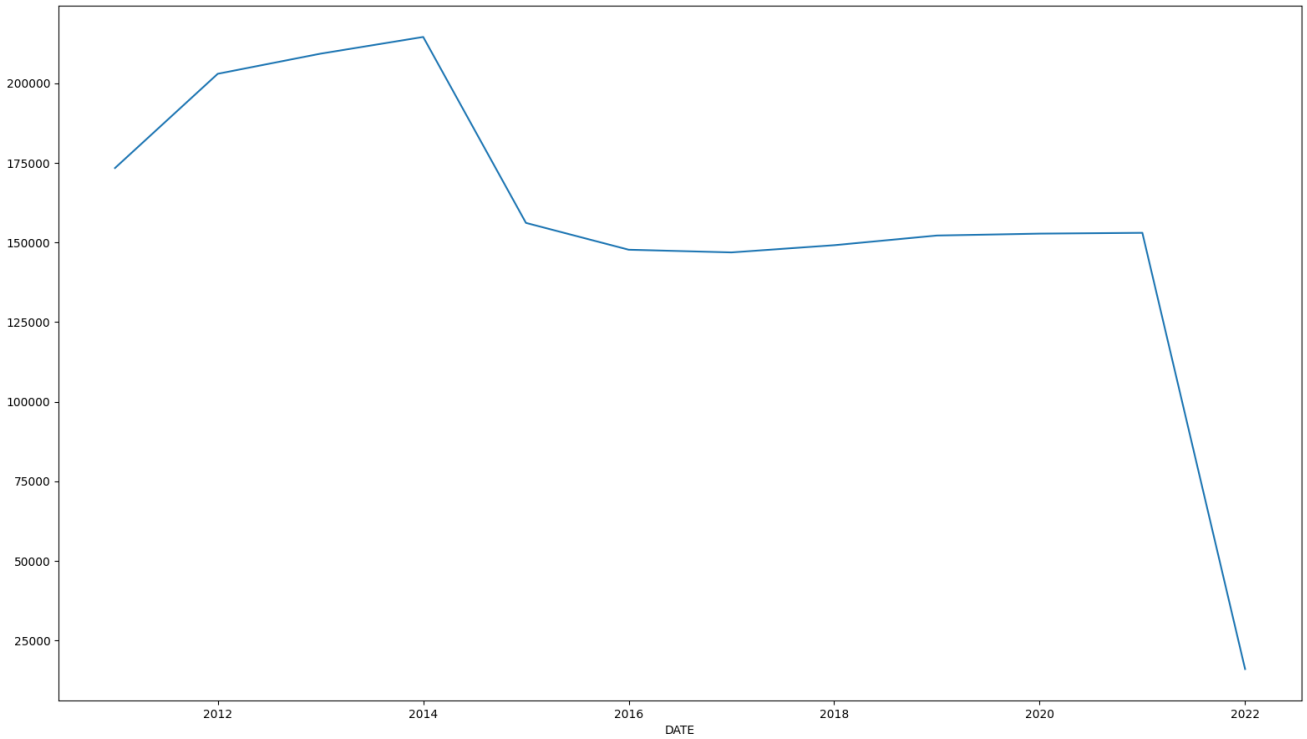
| | DATE | MONTH | TYPE | TOTAL (IN TONS) |
|---|---|---|---|---|
| 1 | 2013-05-31 | May | Curb Garbage | 12594 |
| 2 | 2014-05-31 | May | Curb Garbage | 12375 |
| 3 | 2014-06-30 | June | Curb Garbage | 11770 |
| 4 | 2014-04-30 | April | Curb Garbage | 11564 |
| 5 | 2019-12-31 | December | Yard Waste | 11500 |

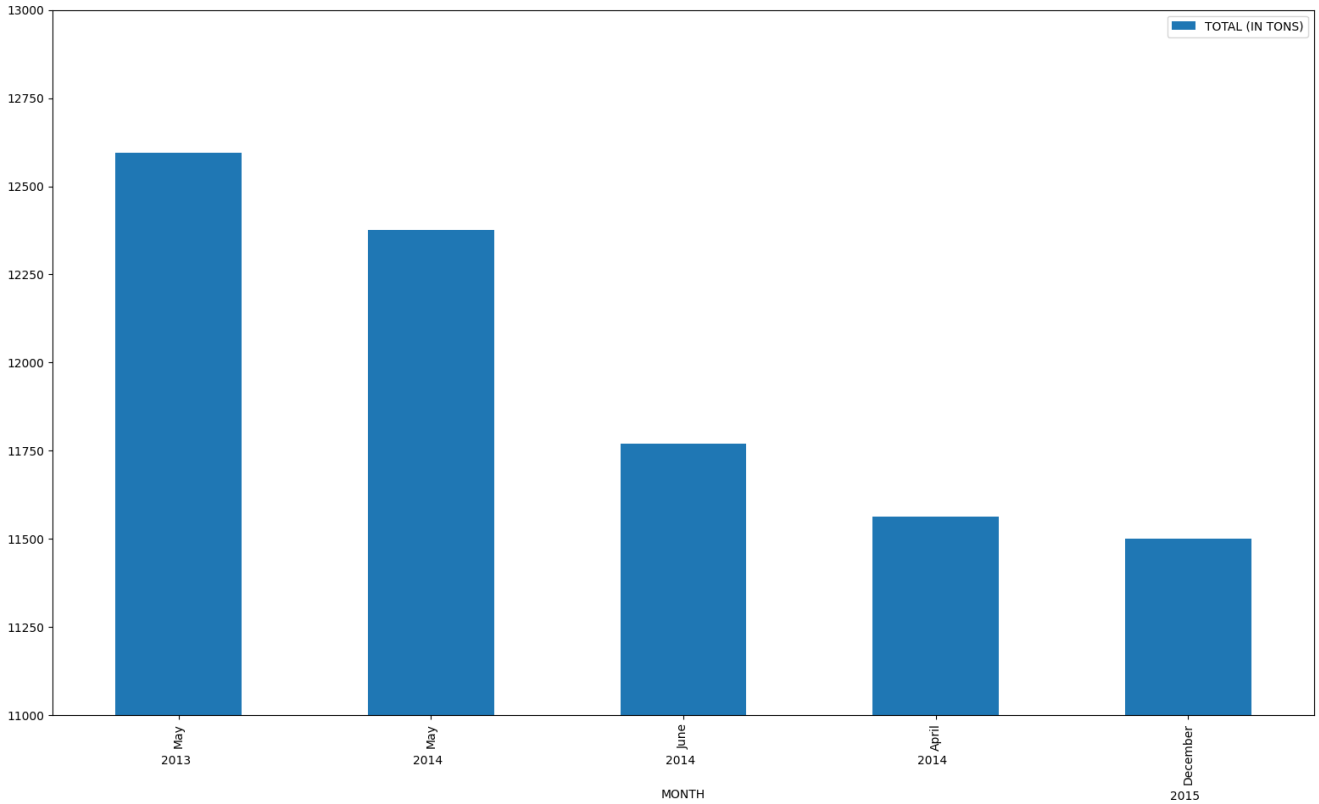
Στη συνέχεια για να κάνουμε τα γραφήματα χρησιμοποιούμε την βιβλιοθήκη matplotlib. Δημιουργούμε 1920 x 1080 εικόνες των γραφημάτων και τις αποθηκέυουμε σε png.

Εδώ συγκεκριμένα στην κατηγοριοποίηση ανά τύπο ανακυκλώσιμων χρειάστηκε να χειριστώ ειδικά τις ακραίες περιπτώσεις (1M τόνοι υλικά και 130 σε αντίθεση με όλους τους άλλους τύπους που κυμαίνονταν μέχρι 200K.) και αποφάσισα απλά να κανω rescale το y axis και να καταγράψω εκείνες τις δύο τιμές «με το χέρι».(βλ. Γραμμές 33-54)

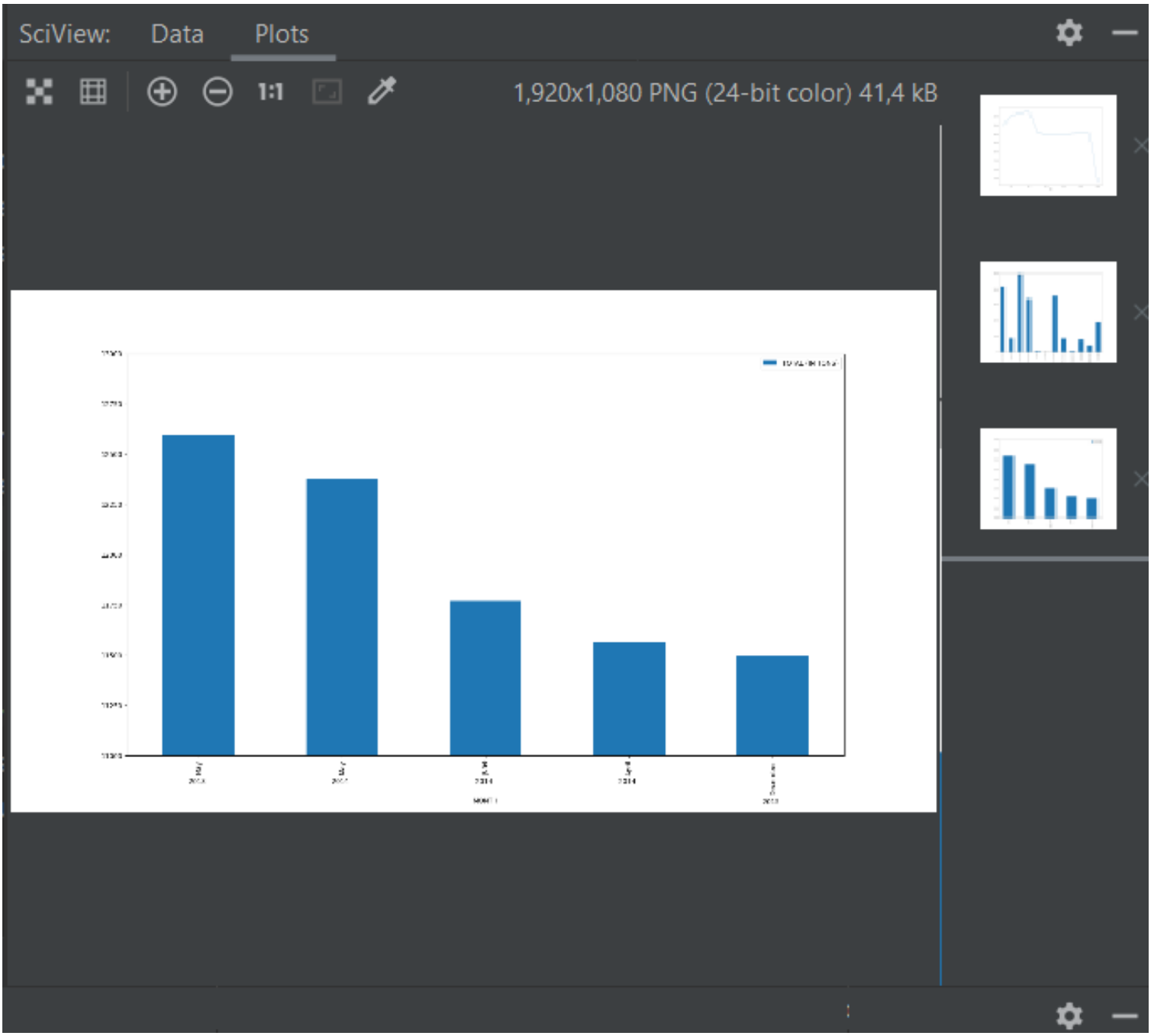
Όμοια και στην κατηγοριοποίηση των μηνών υπήρχε δυσκολία στην εμφάνιση του έτους μαζί με το μήνα στον x axis, έτσι πρόσθεσα τους μήνες με παρόμοιο τρόπο όπως ο παραπάνω.Οι εικόνες των γραφημάτων παρατίθενται παρακάτω.




200000
175000
150000
125000
100000
75000
50000
25000
2012
2014
2016
2018
2020
2022
DATE

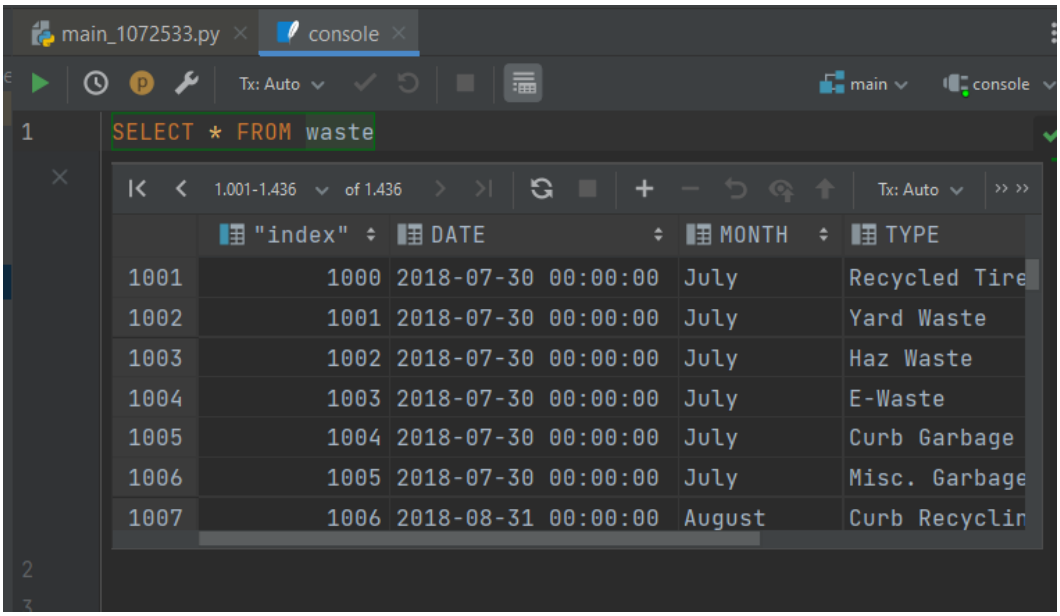

TOTAL (IN TONS)
13000
12750
12500
12250
12000
11750
11500
11250
11000
May 2013
May 2014
June 2014
April 2014
December 2015
MONTH

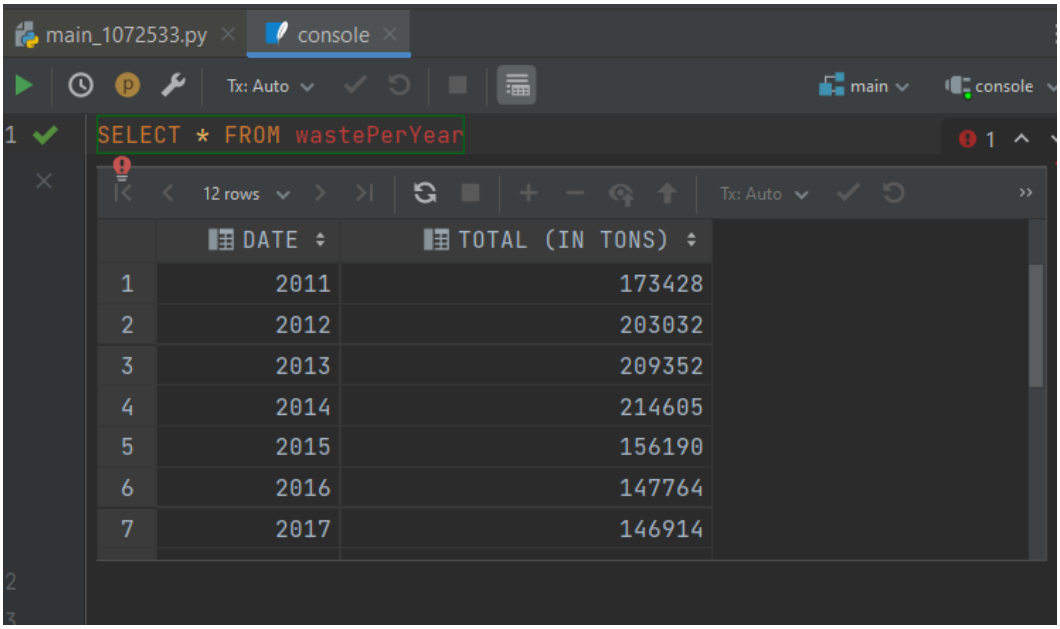
Εμφάνιση των plots με plt.show():

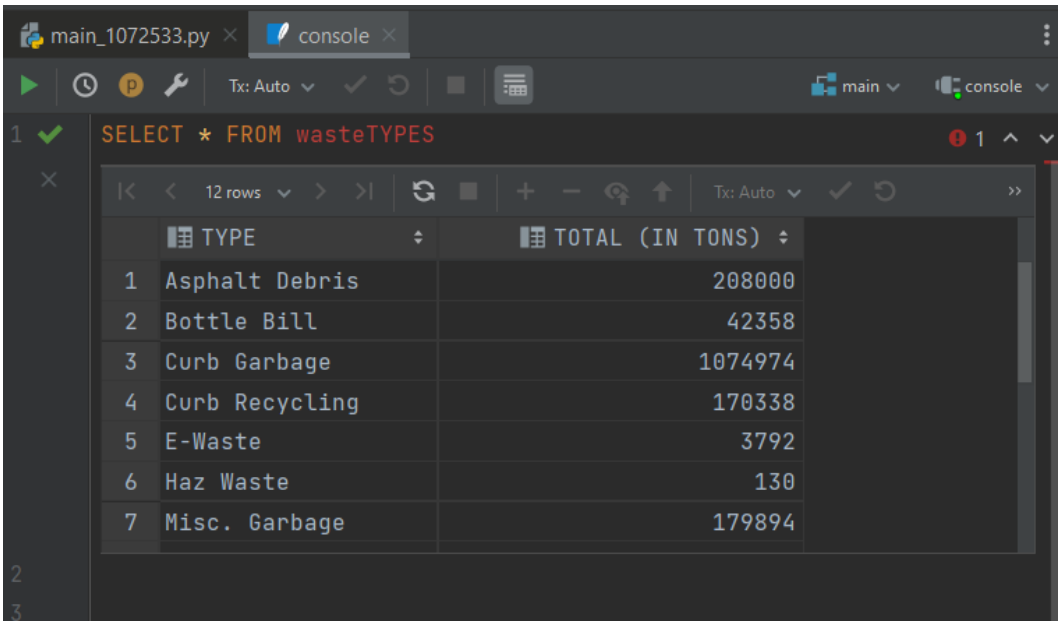


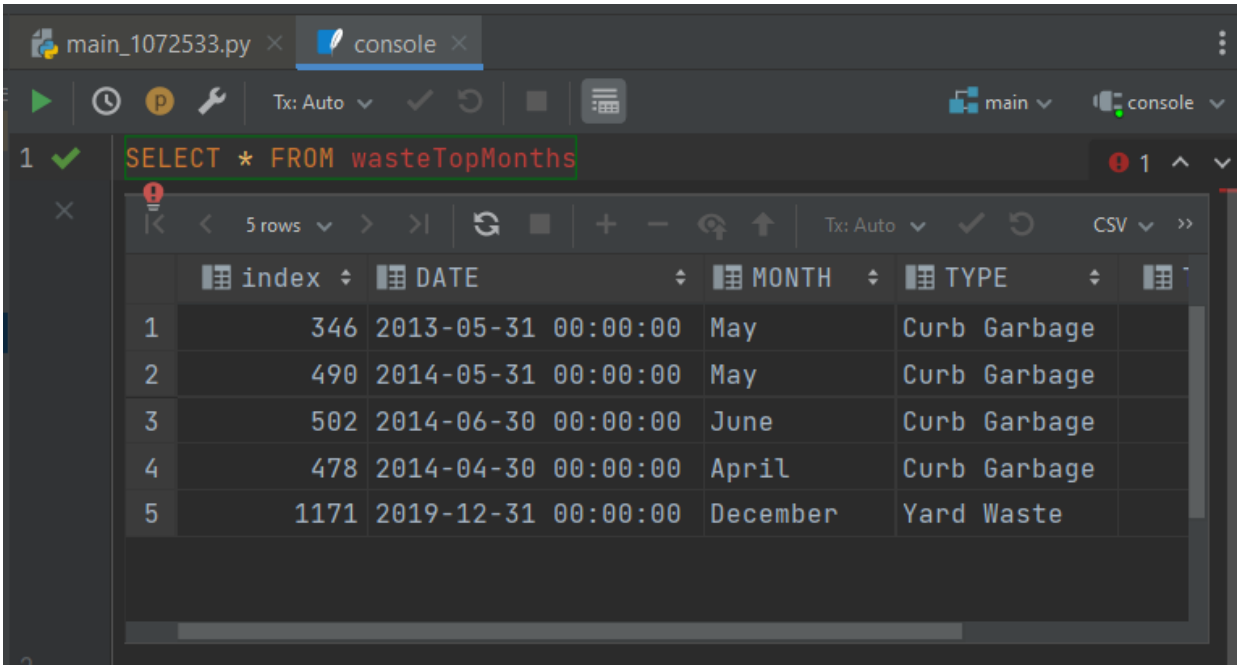

Τέλος, για την ενσωμάτωση σε βάση δεδομένων sql, χρησιμοποιήθηκε η sqlite3,όπου προσφέρει έναν εύκολο και γρήγορο τρόπο για να δημιουργήσει βασεις δεδομένων και να διαχειριστείς sql μέσα στην python.(βλ. 56-61)

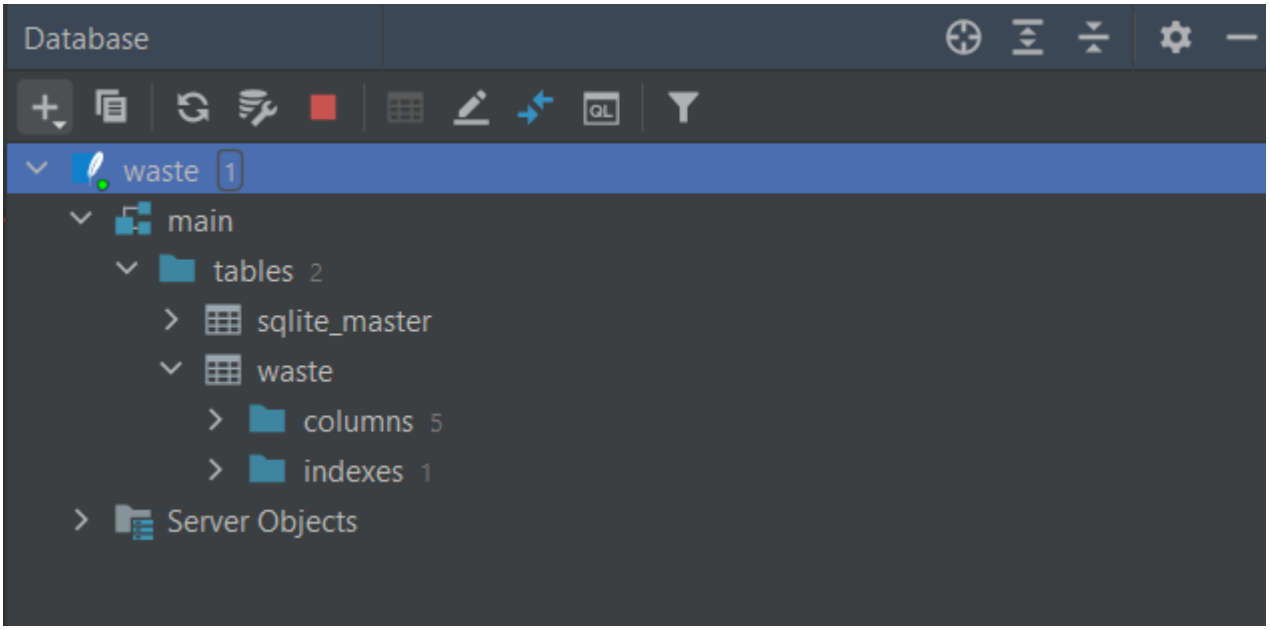
Κάποιες φωτογραφίες και selects:






Το error στα tables δεν ξέρω γιατί το βγάζει, μάλλον γιατι δεν φαίνεται να έχουν δημιουργηθεί στο schema τα κατάλληλα tables. Είναι όμως λειτουργική η βάση δεδομένων ‘waste’.