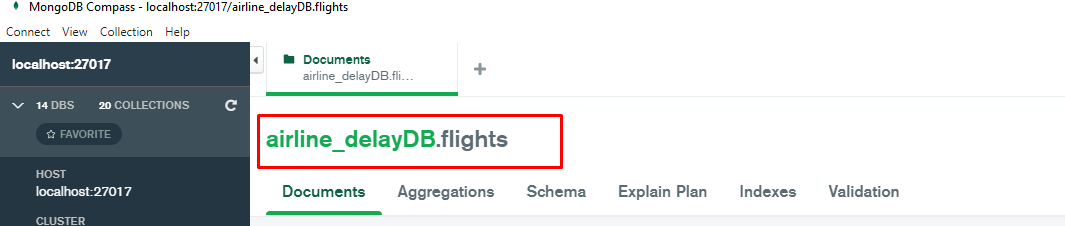
# a) Create collections “flights” inside database “airline\_delayDB”



# b) average arrival delay caused by airlines

all\_docs=collection.aggregate([{'$group':{'\_id':{},'average':{'$avg':'$ARRIVAL\_DELAY'}}},{'$project':{'\_id':0}}])

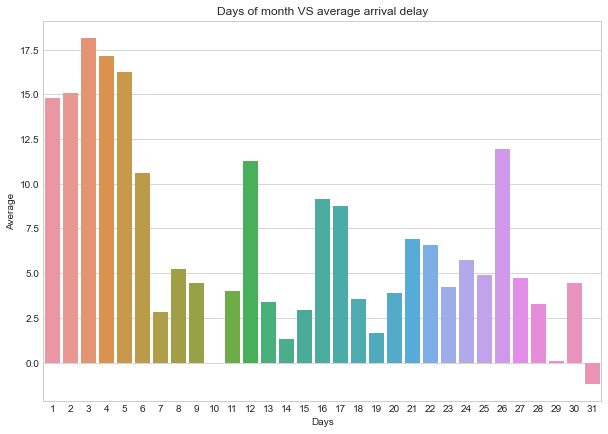
for i in all\_docs:

print(i)

Text

Description automatically generated

c) Days of months with respect to average of arrival delays. [Create a suitable plot using matplotlib/seaborn]



d) Arrange weekdays with respect to the average arrival delays caused. [Create a suitable plot using matplotlib/seaborn]

Chart, bar chart

Description automatically generated

e) Arrange Days of month as per cancellations done in descending order.  [Create a suitable plot using matplotlib/seaborn]

Chart, bar chart, histogram

Description automatically generated

f) Find the busiest airports with respect to day of week. Create a suitable plot using matplotlib/seaborn.

g) Find top 10 Airlines of US. Create a suitable plot using matplotlib/seaborn.

Chart, bar chart

Description automatically generated

h) Finding airlines that make the maximum, minimum number of cancellations.



i) Find and show airlines names in descending that make the most number of diversions made. [Create a suitable plot using matplotlib/seaborn]

Chart, bar chart

Description automatically generated

j) Finding days of month that see the most number of diversion



k) Calculating mean and standard deviation of departure delay for all flights in minutes

Text

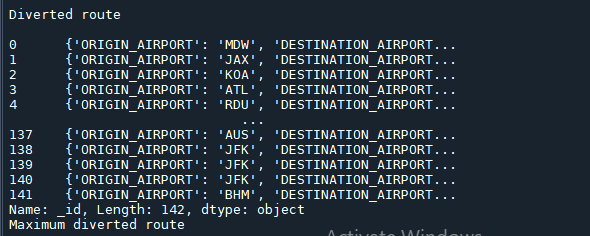
Description automatically generated

l) Calculating mean and standard deviation of arrival delay for all flights in minutes



m) Create a partitioning table “flights\_partition” using partitioned by schema “**CANCELLED”**

n) Finding all diverted Route from a source to destination Airport & which route is the most diverted route



Text

Description automatically generated

o) When is the best time of day/day of week/time of a year to fly with minimum delays ?