**2018-2022**

arr\_flights :Number of flights arrived

arr\_del15 : Number of flights more than 15 mins delayed

carrier\_ct : Number of flights delayed due to air carrier

nas\_ct : Number of flights delayed due to National Aviation System

security\_ct : Number of flights delayed due to security reasons

late\_aircraft\_ct : Number of flights due to aircraft arriving late

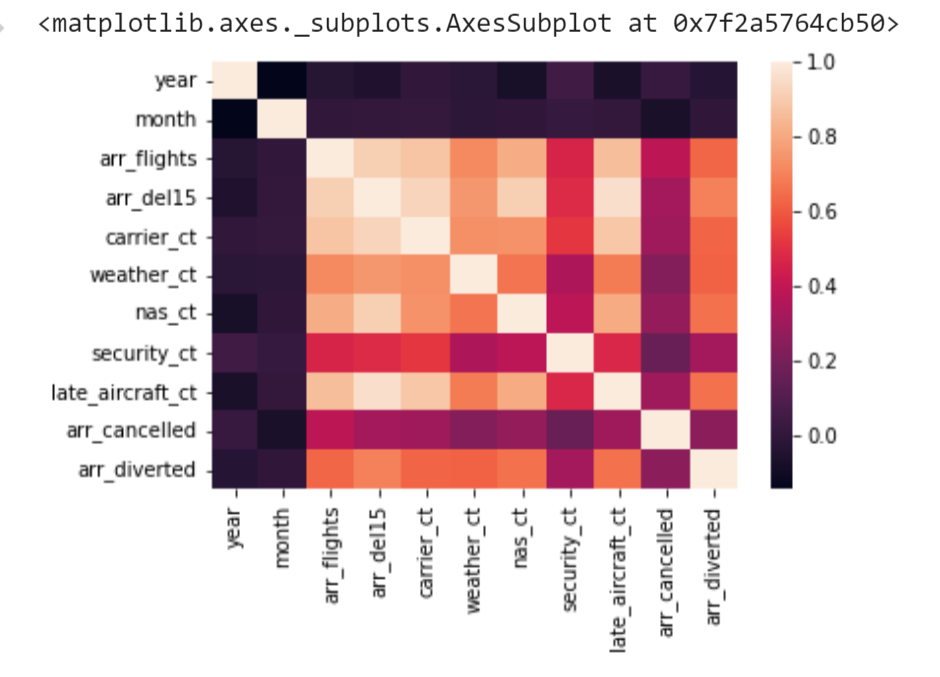
arr\_cancelled : Number of flights canceled

arr\_divert : Number of flights diverted

arr\_delay : Total time (minutes) of delayed flights

carrier\_delay : Total number of carriers delayed

sns.heatmap(number\_of\_flights.corr())

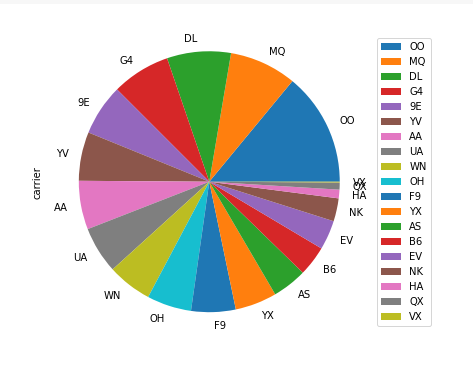


This image shows the correlation between all the variables with each other. We can clearly see that years, and months are not correlated to any other variables, arr\_cancelled and security\_ct have very weak correlation with other variables, and all the other variables are correlated to each other. Which means that all the variables (carrier\_ct, weather\_ct, nas\_ct, late\_aircraft\_ct, arr\_diverted) affect the delay of the flights.

b=df3['carrier'].value\_counts()

b.plot(kind='pie',figsize=(6,6),legend=True)

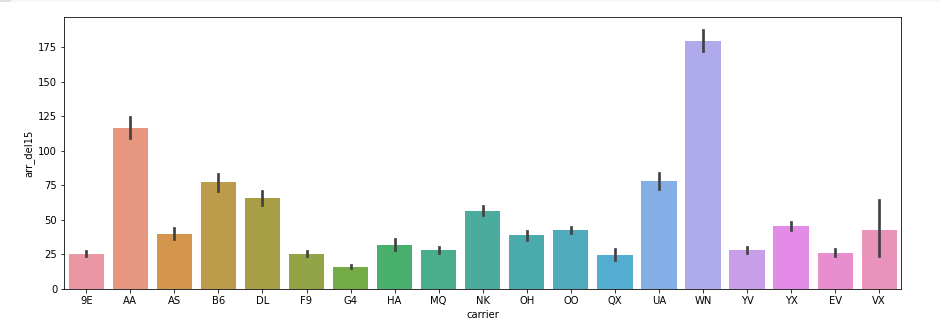
plt.legend(loc=6,bbox\_to\_anchor=(1.0,0.5));

****

Total number of flights carried by carriers

plt.figure(figsize=[15,5])

sb.barplot(data=df3,x='carrier',y='arr\_del15');



Total number of flights Delayed by carriers

fig = px.bar(df4,x='month',y='counts',

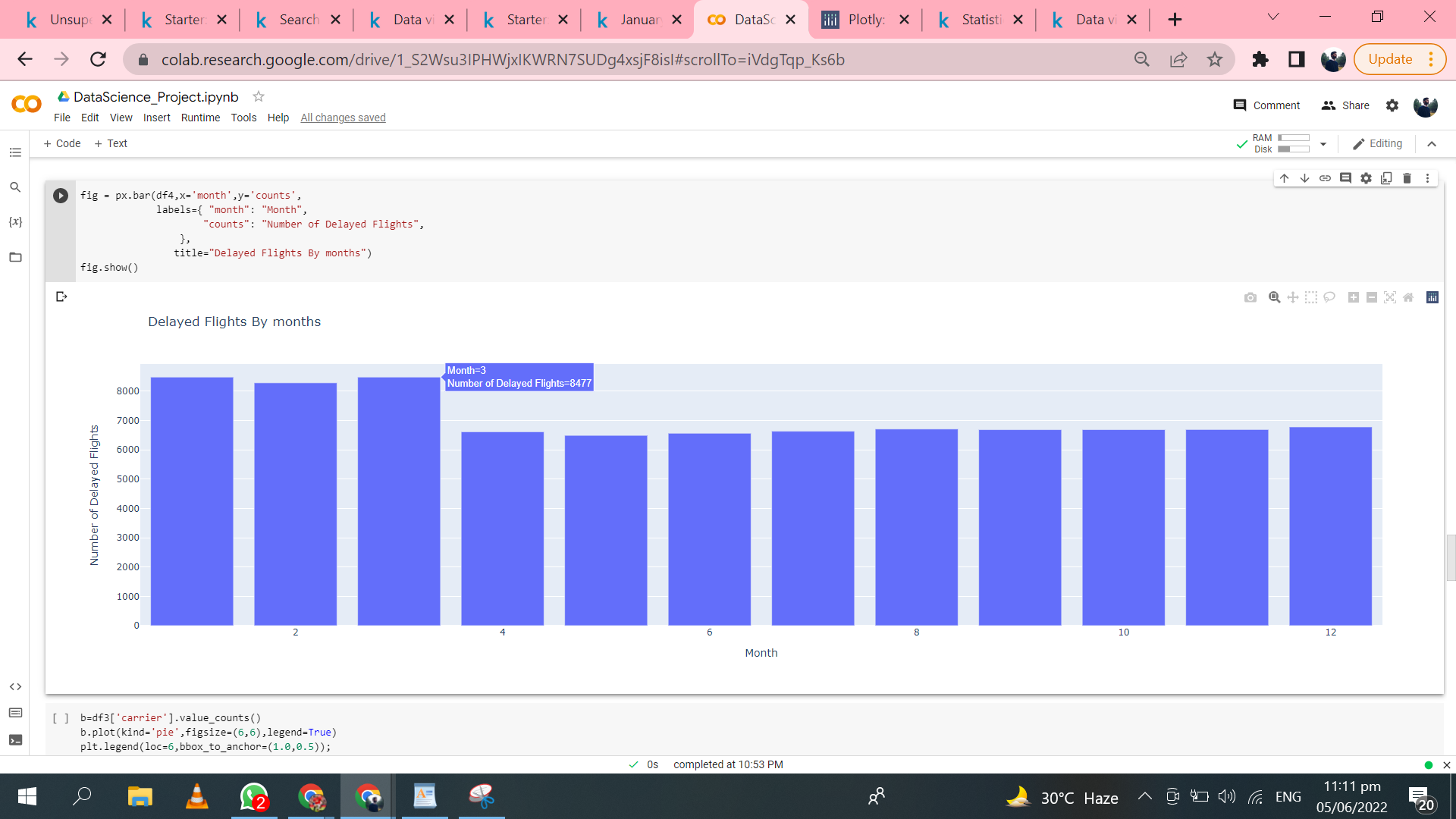
labels={ "month": "Month",

"counts": "Number of Delayed Flights",

},

title="Delayed Flights By months")

fig.show()



Total number of flights Delayed over months

data = df3[['carrier\_ct', 'weather\_ct', 'nas\_ct', 'security\_ct','late\_aircraft\_ct']].mean()

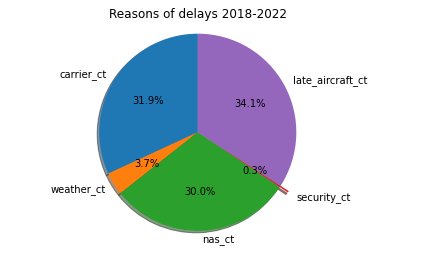
label = ['carrier\_ct', 'weather\_ct', 'nas\_ct', 'security\_ct', 'late\_aircraft\_ct']

plt.pie(data, labels=label, autopct='%1.1f%%', explode=[0,0,0,0.1,0], shadow=True, startangle=90)

plt.title('Reasons of delays 2018-2022')

plt.axis('equal')

plt.show()

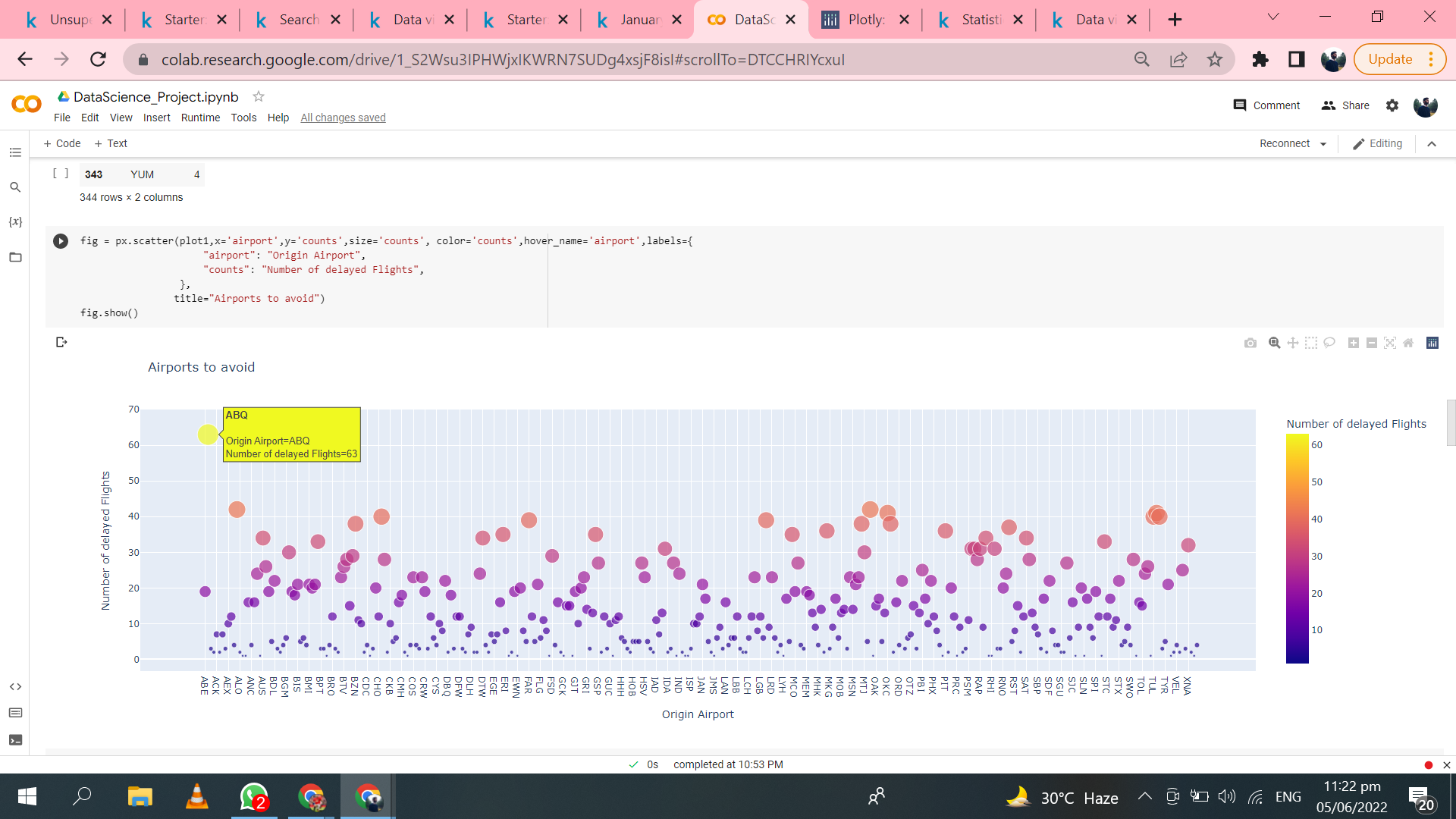


It shows that more of the flight delays are caused by carrier\_ct, late\_aircraft\_ct, nas\_ct.

Weather affects the flight delays but not as much as other variables do.

fig=px.scatter(plot1,x='airport',y='counts',size='counts',color='counts',hover\_name='airport',labels={ "airport": "Origin Airport","counts": "Number of delayed Flights"},title="Airports to avoid")

fig.show()



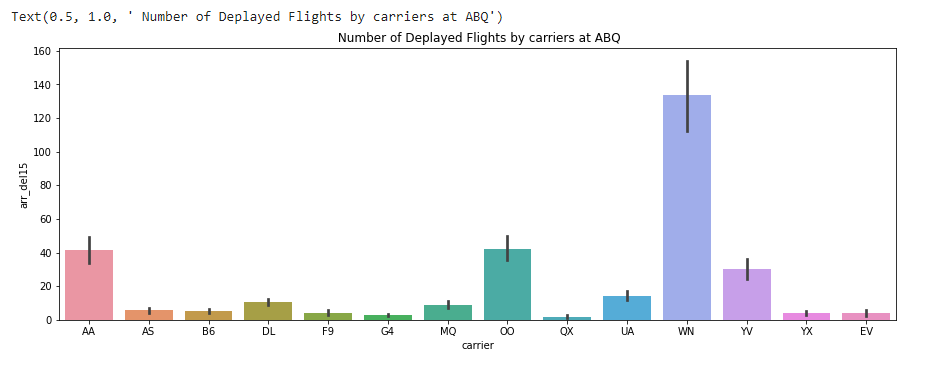
Number of Flights Delayed by for Airports

From the above scatter plot we can confirm that the 3 most busiest airports with a large number of delayed flights are ABQ, ALO, CHO. These are the airports to avoid.

plt.figure(figsize=[15,5])

sb.barplot(data=air,x='carrier',y='arr\_del15');

plt.title(' Number of Delayed Flights by carriers at ABQ')



Total number of flights delayed by carriers at “ABQ” airport

The above bar chart shows that at “ABQ” airport the maximum number of flights are delayed by “WN” carrier.

data = air[['carrier\_ct', 'weather\_ct', 'nas\_ct', 'security\_ct', 'late\_aircraft\_ct']].mean()

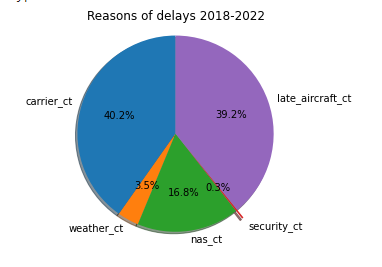
label = ['carrier\_ct', 'weather\_ct', 'nas\_ct', 'security\_ct', 'late\_aircraft\_ct']

plt.pie(data, labels=label, autopct='%1.1f%%', explode=[0,0,0,0.1,0], shadow=True, startangle=90)

plt.title('Reasons of delays 2018-2022')

plt.axis('equal')

plt.show()

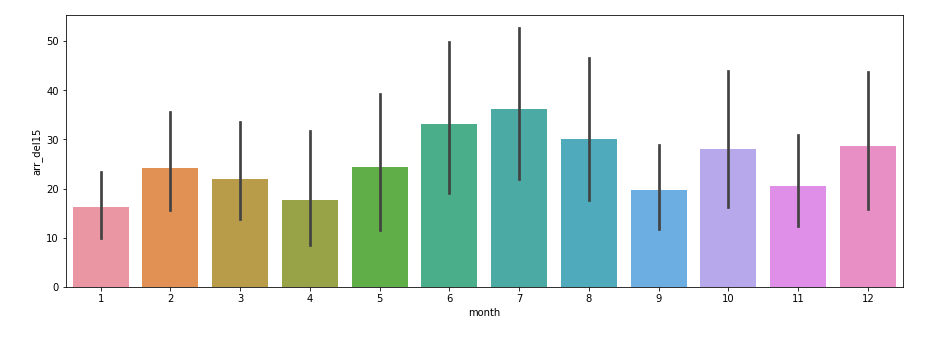


Causes of flights delayed at “ABQ” airport

The above pie chart shows that more of the flight delays are caused by carrier\_ct, late\_aircraft\_ct, and delays because of “nas\_ct” are less than overall delays. And delays due to weather are way less than other variables.

plt.figure(figsize=[15,5])

sb.barplot(data=air,x='month',y='arr\_del15');

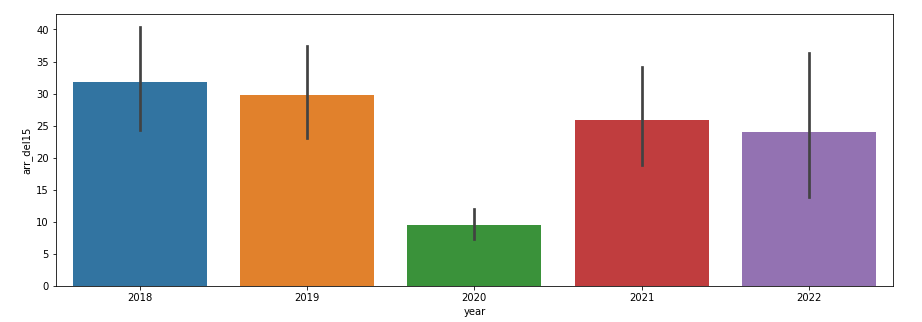


Total number of flights delayed at “ABQ” airport

If we see the pattern of flights delayed over months then it’s clear more of the flights are delayed in the summers.

plt.figure(figsize=[15,5])

sb.barplot(data=air,x='year',y='arr\_del15');



Total number of flights delayed at “ABQ” airport

From the above bar chart we can find the pattern that over the 5 years flight delays have decreased. In addition to that there was a global pandemic starting after 2019 and there were restrictions on international flights. That’s why the above plot is showing the decrease in flight delays after 2019.