In [2]:

```
#S04 T01: Visualització gràfica d'un dataset
```

In [3]:

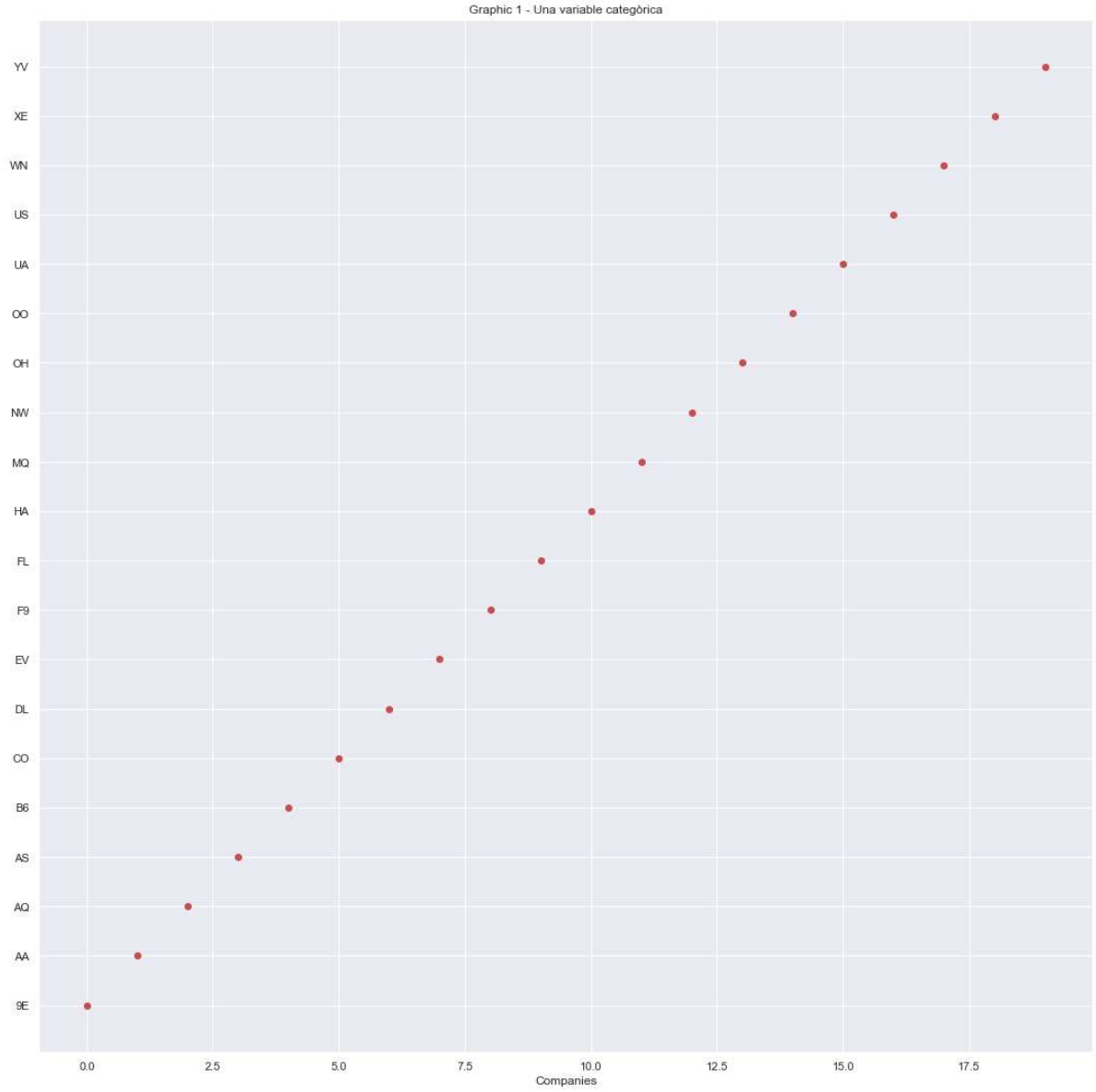
```
#Nivell 1
```

In [4]:

```
#Exercici 1
#Resumeix gràficament el data set DelayedFlights.csv
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

airlines_df = pd.read_csv('Python/DelayedFlights.csv')
```

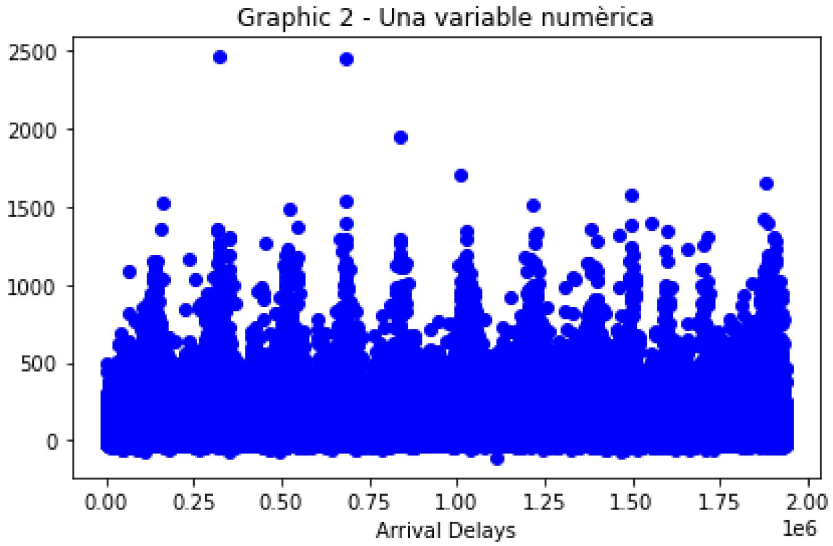
In [12]:

```
#Exercici 1.a
import matplotlib.pyplot as plt01

#Crea almenys una visualització per:

#Graphic 1: Una variable categòrica (UniqueCarrier)

x = np.unique(airlines_df['UniqueCarrier'])
x = np.sort(x)

plt01.title("Graphic 1 - Una variable categòrica")
plt01.xlabel("Companies")
plt01.plot(x,'ro')
plt01.tight_layout()
plt01.show()
```



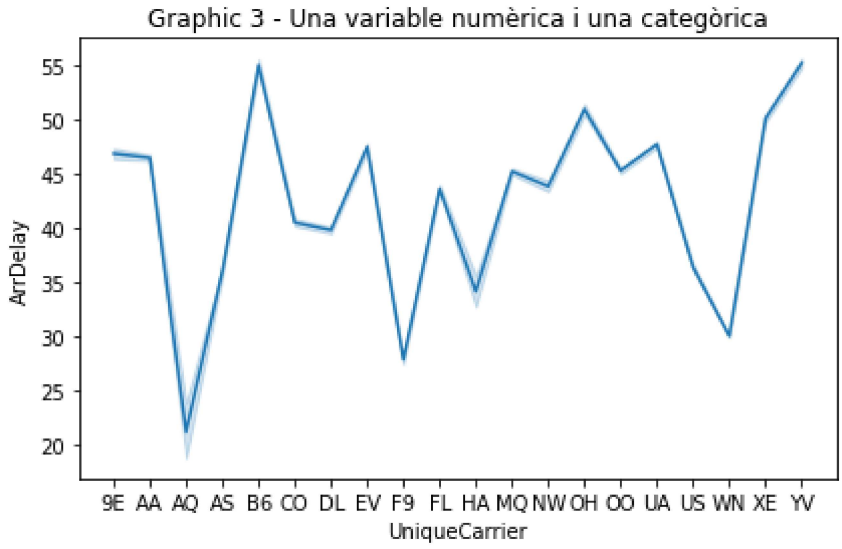
In [6]:

```
#Exercici 1.b
import matplotlib.pyplot as plt02

#Graphic 2: Una variable numèrica (ArrDelay)

x = np.array(airlines_df[['ArrDelay']])
plt02.title("Graphic 2 - Una variable numèrica")
plt02.xlabel("Arrival Delays")
plt02.plot(x,'bo')
plt02.tight_layout()
plt02.show()
```



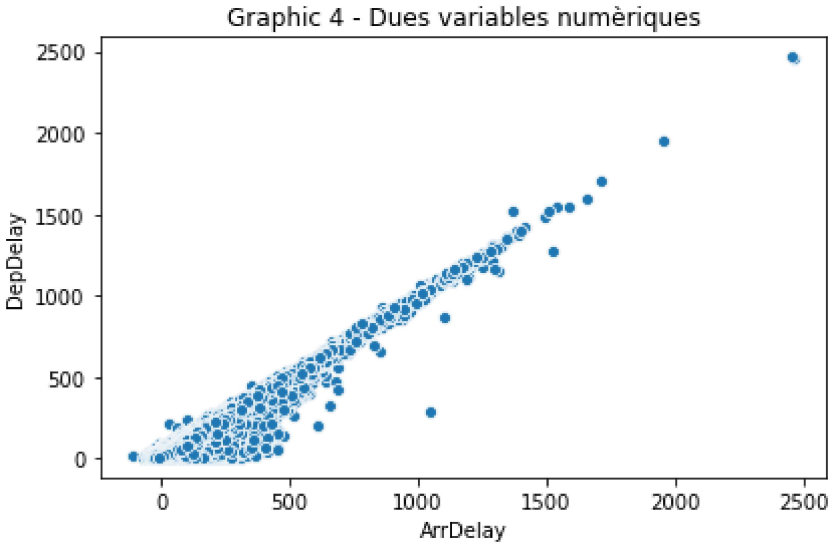
In [7]:

```
#Exercici 1.c
import matplotlib.pyplot as plt03

#Graphic 3: Una variable numèrica i una categòrica (ArrDelay i UniqueCarrier)
#import seaborn as sns

plt03.title("Graphic 3 - Una variable numèrica i una categòrica")
plt03.xlabel("Companies")
plt03.ylabel("Arrival Delays")
sns.lineplot(x="UniqueCarrier", y="ArrDelay", data=airlines_df, palette ="pastel")
plt03.tight_layout()
plt03.show()
```



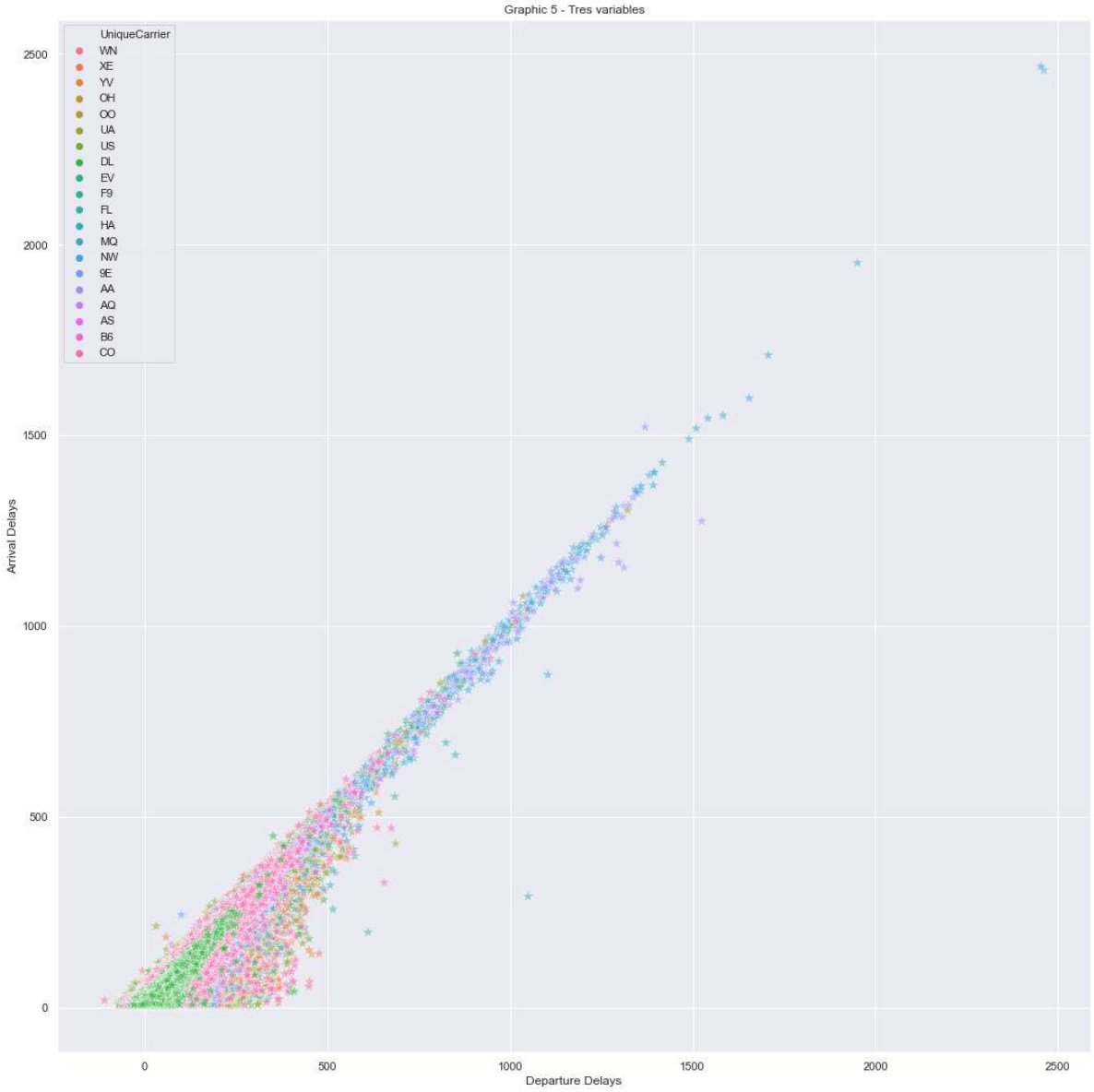
In [8]:

```
#Exercici 1.d
import matplotlib.pyplot as plt04
import seaborn as sns

#Graphic 4: Dues variables numèriques (ArrDelay i DepDelay)
plt04.title("Graphic 4 - Dues variables numèriques")
plt04.xlabel("Departure Delays")
plt04.ylabel("Arrival Delays")
sns.scatterplot(data=airlines_df, x="ArrDelay", y="DepDelay")
plt04.tight_layout()
plt04.show()
```



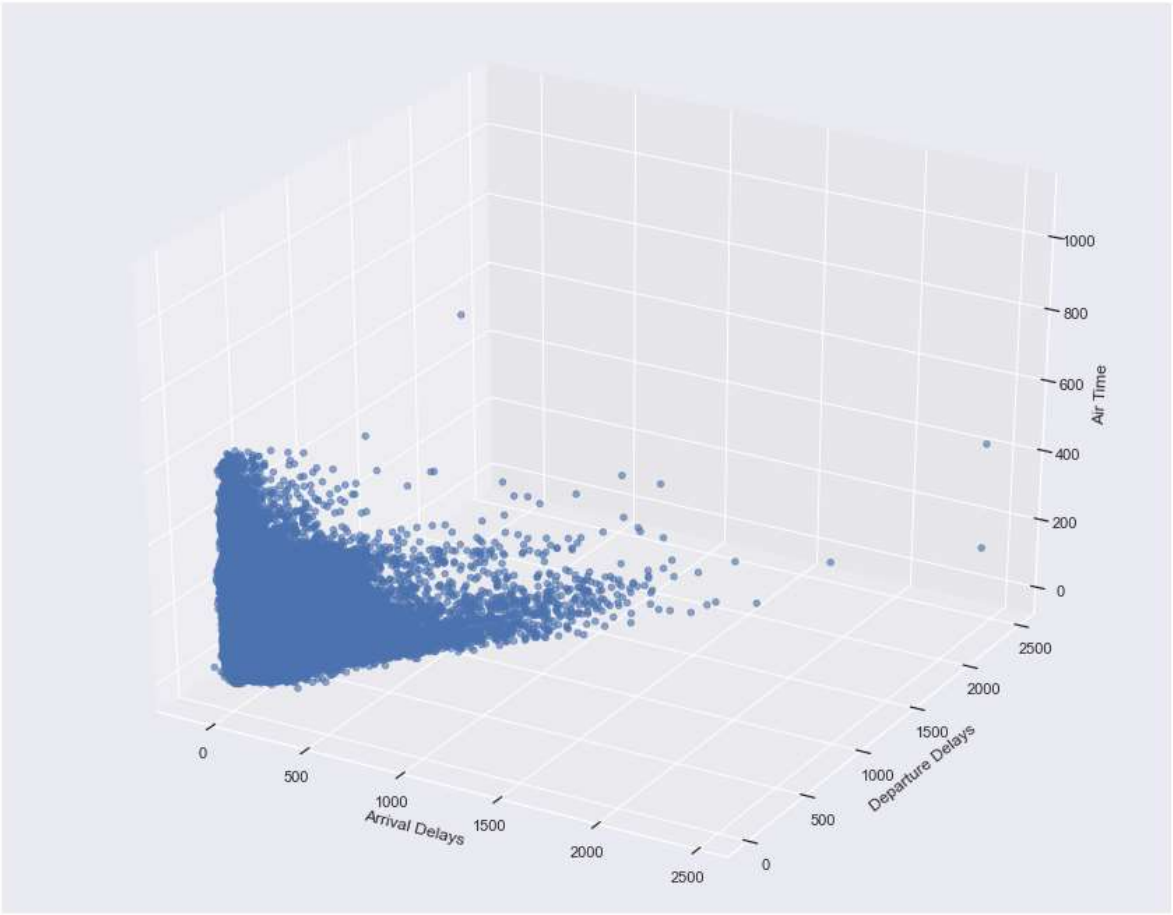
In [9]:

```
#Exercici 1.e
import matplotlib.pyplot as plt05
import seaborn as sns

#Graphic 5: Tres variables (ArrDelay, DepDelay i UniqueCarrier)

sns.set(rc={"figure.figsize":(15, 15)})
sns.scatterplot(data=airlines_df, x="ArrDelay", y="DepDelay", color = 'red',
                marker = '*', hue = "UniqueCarrier", alpha=.5, s= 150).set(title='Graphic 5

plt05.xlabel("Departure Delays")
plt05.ylabel("Arrival Delays")
plt05.legend(loc='best')
plt05.tight_layout()
plt05.show()
```



In [10]:

```
#Exercici 1.e
import matplotlib.pyplot as plt06

#Graphic 6: Més de tres variables (ArrDelay, DepDelay, AirTime i UniqueCarrier).

fig = plt06.figure(figsize=(15,12))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(airlines_df['ArrDelay'],
           airlines_df['DepDelay'],
           airlines_df['AirTime'],alpha=.6)

ax.set_xlabel('Arrival Delays')
ax.set_ylabel('Departure Delays')
ax.set_zlabel('Air Time')
plt06.show()
```



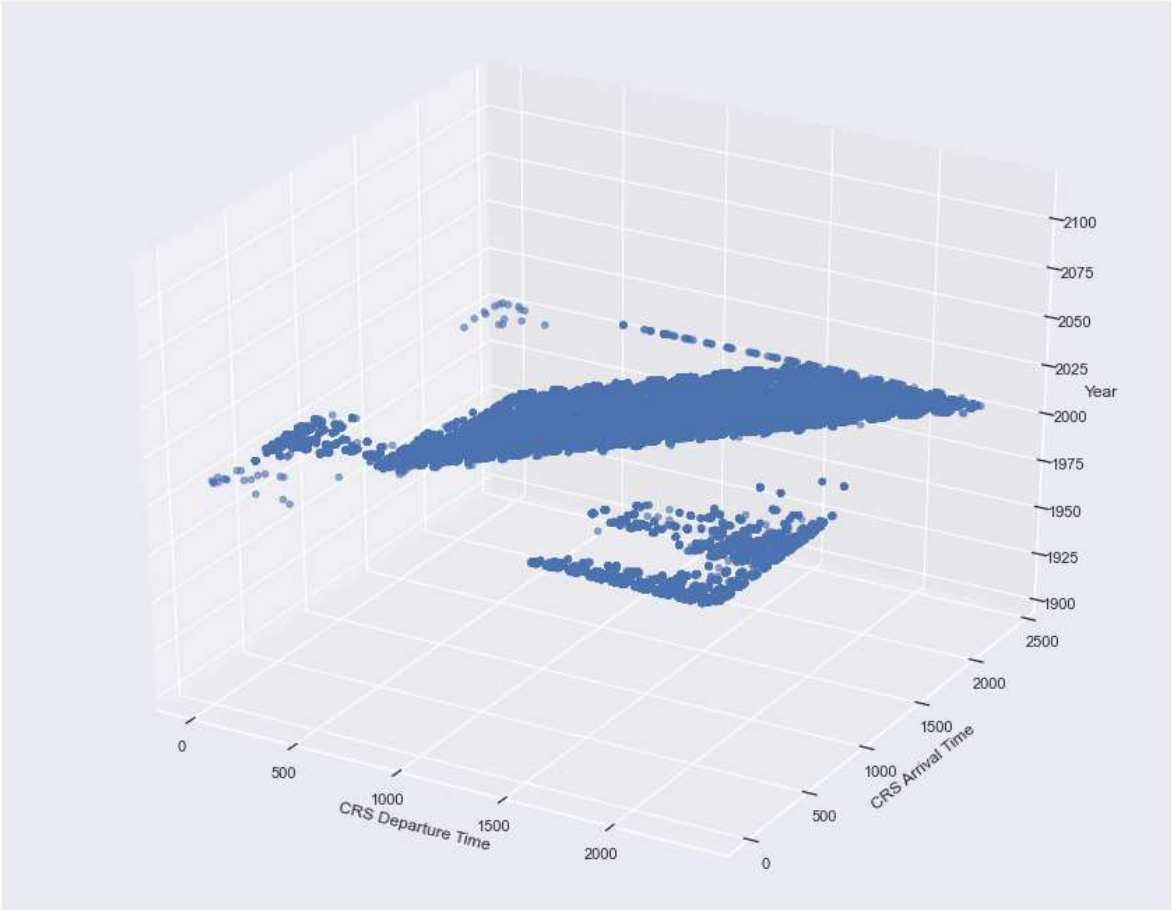
In [14]:

```
#Exercici 2
#Exporta els gràfics com imatges o com html.

plt01.savefig('Python/air_df01.png')
plt02.savefig('Python/air_df02.png')
plt03.savefig('Python/air_df03.png')
plt04.savefig('Python/air_df04.png')
plt05.savefig('Python/air_df05.png')
plt06.savefig('Python/air_df06.png')
```

```
<Figure size 1080x1080 with 0 Axes>
```

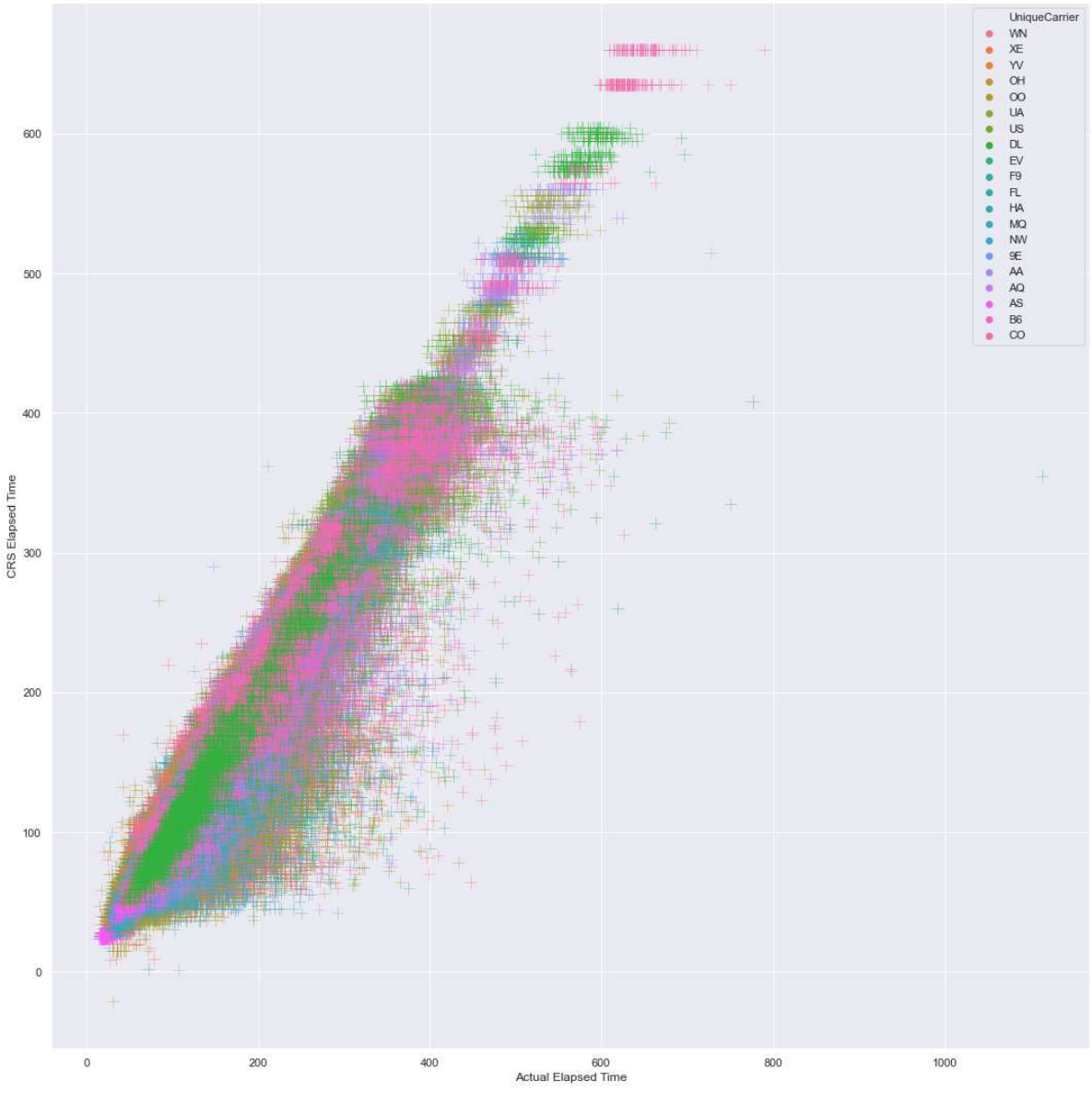
In [17]:

```
#Exercici 3.1
#Integra les visualitzacions gràfiques, en la tasca 5, del Sprint 3.
from matplotlib import pyplot as plt

airlines_df[["Year","DepTime","CRSDepTime","ArrTime","CRSArrTime"]]

fig = plt.figure(figsize=(15,12))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(airlines_df['CRSDepTime'],
           airlines_df['CRSArrTime'],
           airlines_df['Year'],alpha=.6)

ax.set_xlabel('CRS Departure Time')
ax.set_ylabel('CRS Arrival Time')
ax.set_zlabel('Year')
plt.show()
```



In [19]:

```
#Exercici 3.2
#Integra les visualitzacions gràfiques, en la tasca 5, del Sprint 3.
from matplotlib import pyplot as plt
import seaborn as sns

airlines_df[["ActualElapsedTime","CRSElapsedTime","UniqueCarrier","TaxiIn","TaxiOut"]]

sns.set(rc={"figure.figsize":(15, 15)})
sns.scatterplot(data=airlines_df, x="ActualElapsedTime", y="CRSElapsedTime", color = 'green
                marker = '+', hue = "UniqueCarrier", alpha=.5, s= 150)

plt05.xlabel("Actual Elapsed Time")
plt05.ylabel("CRS Elapsed Time")
plt05.legend(loc='best')
plt05.tight_layout()
plt05.show()
```



In [ ]: