Qraded Quiz- IBM Applied Data Science Capstone

Week 1-Graded Quiz: Data Collection API and Web Scraping

1. After you performed a **GET** request on the Space X API and convert the response to a dataframe using **pd.json_normalize**. What year is located in the first row in the column static_fire_date_utc?

2006

2. Using the **API**, how many Falcon 9 launches are there after we remove Falcon 1 launches?

90

3. At the end of the **API** data collection process, how many missing values are there for the column landingPad?

26

4. After making a request to the Falcon9 Launch Wiki page and creating a BeautifulSoup object what is the output of: **soup.title**

- < td colspan="9" > First flight of Falcon 9 v1.0. < sup class="reference" id="cite_ref-sfn20100604_17-0"
- <table class="wikitable plainrowheaders collapsible" style="width: 100%;">
- **<title> List of Falcon 9 and Falcon Heavy launches - Wikipedia </title>**

Week 1-Graded Quiz: Data Wrangling

1. How many launches came from CCAFS SLC 40?

55

```
df['LaunchSite'].value_counts()
CCAFS SLC 40     55
KSC LC 39A       22
VAFB SLC 4E      13
```

2. What was the success rate?

- 80%
- 40%
- **67%**

```
df['Class']
```

	Class
0	0
1	0
2	0
89	1

```
df['Class'].mean()
0.66
```

3. In the lab you used the method .value_counts() to determine the number and occurrence of each orbit in the column Orbit. What was the value for Orbit with the column name GTO?

27

```
df['Orbit'].value_counts()
GTO       27
ISS       21
VLEO      14
PO         9
LEO        7
SSO        5
MEO        3
ES-L1      1
HEO        1
SO         1
GEO        1
```

4. How many landing outcomes in the column landing_outcomes had a value of none?

19

```
landing_outcomes = df['Outcome'].value_counts()
landing_outcomes
True ASDS       41
```

```
None None      19
True RTLS      14
False ASDS      6
True Ocean      5
False Ocean     2
None ASDS       2
False RTLS      1
```

Week 2-Graded Quiz: Exploratory Data Analysis using SQL

1. Which of the following will retrieve up to 20 records from the spacex table?

- **SELECT * from SPACEXTBL LIMIT 20**
- SELECT * from SPACEXTBL MAX 20
- SELECT * from SPACEXTBL where count(*)=20
- SELECT TOP 20 rows from SPACEXTB

2. Which of the following queries display the minimum payload mass?

- select payload_mass__kg_ from SPACEXTBL order by payload_mass__kg_ group by booster_version LIMIT 1
- select payload_mass__kg_ from SPACEXTBL order by payload_mass__kg_ desc LIMIT 1
- **select min(payload_mass__kg_) from SPACEXTBL**
- select payload_mass__kg_ from SPACEXTBL where payload_mass__kg_=(select max(payload_mass__kg_) from SPACEXTBL) LIMIT 1

3. You are writing a query that will give you the total payload_mass_kg carried by the booster versions. The mass should be stored in the mass column. You want the result column to be called "Total_Payload_Mass". Which of the following SQL queries is correct?

- SELECT count(PAYLOAD_MASS__KG_) as Total_Payload_Mass from SPACEXTBL
- **SELECT sum(PAYLOAD_MASS__KG_) as Total_Payload_Mass from SPACEXTBL**
- SELECT sum(PAYLOAD_MASS__KG_) from SPACEXTBL

In SQL, **Sum** is used for numerical variables**. Count** is used for categorical variables.

4. Which of the following query is used to display the mission outcome **counts** for each launch site?

- **select count("Mission_Outcome") as MISSION_OUTCOME_COUNT,Launch_Site from SPACEXTBL group by "Launch_Site";**
- select sum("Mission_Outcome") as MISSION_OUTCOME_COUNT,Launch_Site from SPACEXTBL group by "Launch_Site";

5. What are the unique launch sites mentioned in the Spacex table?

- **CCAFS LC-40,KSC LC-39A, VAFB SLC-4E , CCAFS SLC-40**
- CCAFS LC-40,KSC LC-39B,VAFB SLC-4k , CCAFS SLC-40
- None of the Above
- CCAS LC-40,KSC LC-39A,VAFB SLC-4E , CCAFS SLC-80

```
%sql SELECT DISTINCT Launch_Site FROM SPACEXTABLE;
Launch_Site
CCAFS LC-40
VAFB SLC-4E
KSC LC-39A
CCAFS SLC-40
```

Week 2-Graded Quiz: Exploratory Data Analysis using Pandas and Matplotlib

1. What type of data does **a Bar Chart** best represent?

- Location Data
- Numerical
- **Categorical**
- None of the above

2. What are the total number of columns in the features dataframe after applying one hot encoding to columns Orbits, LaunchSite, LandingPad and Serial. Here the **features dataframe** consists of the following columns FlightNumber', 'PayloadMass', 'Orbit', 'LaunchSite', 'Flights', 'GridFins', 'Reused', 'Legs', 'LandingPad', 'Block', 'ReusedCount', 'Serial'

- 120
- **80**
- 83
- 96

3. The **catplot** code to show the scatterplot of FlightNumber vs LaunchSite with x as FlightNumber, and y to Launch Site and hue to 'Class' is

- sns.catplot(y="LaunchSite",x="FlightNumber",hue="Class", data=df, aspect = 1,kind='cat')
  plt.ylabel("Launch Site",fontsize=15)
  plt.xlabel("Flight Number",fontsize=15)
  plt.show()

- **sns.catplot(y="LaunchSite",x="FlightNumber",hue="Class", data=df, aspect = 1)**
  plt.ylabel("Launch Site",fontsize=15)
  plt.xlabel("Flight Number",fontsize=15)
  plt.show()
- sns.catplot(y="LaunchSite",x="FlightNumber",hue="Class", data=df, aspect = 1,kind='scatter')
  plt.ylabel("Launch Site",fontsize=15)
  plt.xlabel("Flight Number",fontsize=15)
  plt.show()
- sns.catplot(y="LaunchSite",x="FlightNumber",hue="Class", col="Class", data=df, aspect = 1)
  plt.ylabel("Launch Site",fontsize=15)
  plt.xlabel("Flight Number",fontsize=15)
  plt.show()

Week 3-Graded Quiz: Interactive Visual Analytics and Dashboard

1. How can you add marking objects such as circles, markers, or lines on a Folium map? (Click all choices that apply)
- **map.add_child(object)**
- add_node(map, object)
- map.add_to(object)
- **object.add_to(map)**

2. If you want to add multiple markers with similar coordinates on the Folium map, which Folium plugin you should use?
- **MarkerCluster**
- MarkerGroup
- MarkerContainer
- Markers should be add to map directly without any extra layer

3. Which attribute is used to provide available selections (such as a list of launch sites) for a Plotly DropDown input?
- **options**
- values
- input
- placeholder

4. How can we associate the result of a callback function (like a Ploty figure) to an element defined in the application layout?
- Using component name
- Dash automatically render the result of a callback function
- **Using a unique component id**

5. Can we add multiple input components to a dash callback function?
- **Yes**
- No

Week 4-Graded Quiz: Predictive Analysis

1. How many records were there in the test sample?

18

2. For Support Vector Machines, what kernel has the best result on the validation dataset.
- rbf
- **sigmoid**
- linear

3. After selecting the best hyperparameters for the decision tree classifier using the validation data, what was the accuracy achieved on the test data?
- **83.33%**
- 73.33%
- 93.33%