IBM_Data_Science_Professional_Specialization

Rex Coleman

2024-11-13

IBM Data Science Professional Certificate Portfolio

Overview

This repository showcases my journey through the IBM Data Science Professional Certificate program. This 10 course specialization provides a comprehensive overview of data science fundamentals, emphasizing practical applications and real-world scenarios. Each project aligns with the skills and knowledge necessary for a data scientist, culminating in this capstone project forecasting SpaceX rocket launch success: report, code. Specifically for this repository, I showcase 4 skills that are as clearly demonstrated in other parts of my portfolio: i) web scraping, ii) SQL, iii) interactive map plotting with Folium and iv) dashboard plotting with Plotly and Dash.

Certificate Highlights

- **Comprehensive Skill Set:** Master the practical skills used by data scientists, including Python, SQL, data visualization, and machine learning.
- **Real-World Applications:** Implement data science techniques to solve practical problems in various domains.
- **Professional Tools:** Work with essential tools like Jupyter notebooks, GitHub, RStudio, and IBM Watson Studio.
- **Portfolio Projects:** Develop a diverse portfolio that showcases data extraction, analysis, visualization, and machine learning models.

Courses and Key Projects

Course 1: What is Data Science?

Overview: Understand the fundamentals of data science and its significance. **Key Project:** Define data science, explore career paths, and gather insights from seasoned professionals.

Course 2: Tools for Data Science

Overview: Get acquainted with the essential tools in a data scientist's toolkit. **Key Project:** Utilize tools like Python, R, SQL, Jupyter notebooks, and GitHub.

Course 3: Data Science Methodology

Overview: Learn the structured approach to tackling data science problems. **Key Project:** Apply the CRISP-DM methodology to analyze case studies.

Course 4: Python for Data Science, AI & Development

Overview: Gain proficiency in Python programming for data science. **Key Project:** Develop Python code using libraries like Pandas and Numpy.

2020 [edit]

In late 2019, Gwynne Shotwell stated that SpaceX hoped for as many as 24 launches for Starlink satellites in 2020,[490] in addition to 14 or 15 non-Starlink launches. At 26 launches, 13 of which for Starlink satellites, Falcon 9 had its most prolific year, and Falcon rockets were second most prolific rocket family of 2020, only behind China's Long March rocket family.[491]

[hide] Flight No.	Date and time (UTC)	Version, Booster[b]	Launch site	Payload[c]	Payload mass	Orbit	Customer	Launch outcome	Booster landing
78	7 January 2020, 02:19:21[492]	F9 B5 △ B1049.4	CCAFS, SLC-40	Starlink 2 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)
	Third large batch and second operational flight of Starlink constellation. One of the 60 satellites included a test coating to make the satellite less reflective, and thus less likely to interfere with ground-based astronomical observations.[493]								
79	19 January 2020, 15:30[494]	F9 B5 △ B1046.4	KSC, LC-39A	Crew Dragon in-flight abort test[495] (Dragon C205.1)	12,050 kg (26,570 lb)	Sub-orbital[496]	NASA (CTS)[497]	Success	No attempt
	An atmospheric test of the Dragon 2 abort system after Max Q. The capsule fired its SuperDraco engines, reached an apogee of 40 km (25 mi), deployed parachutes after reentry, and splashed down in the ocean 31 km (19 mi) downrange from the launch site. The test was previously slated to be accomplished with the Crew Dragon Demo-1 capsule;[498] but that test article exploded during a ground test of SuperDraco engines on 20 April 2019.[419] The abort test used the capsule originally intended for the first crewed flight.[499] As expected, the booster was destroyed by aerodynamic forces after the capsule aborted.[500] First flight of a Falcon 9 with only one functional stage — the second stage had a mass simulator in place of its engine.								
80	29 January 2020, 14:07[501]	F9 B5 △ B1051.3	CCAFS, SLC-40	Starlink 3 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)
	Third operational and fourth large batch of Starlink satellites, deployed in a circular 290 km (180 mi) orbit. One of the fairing halves was caught, while the other was fished out of the ocean.[502]								
81	17 February 2020, 15:05[503]	F9 B5 △ B1056.4	CCAFS, SLC-40	Starlink 4 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Failure (drone ship)
	Fourth operational and fifth large batch of Starlink satellites. Used a new flight profile which deployed into a 212 km × 386 km (132 mi × 240 mi) elliptical orbit instead of launching into a circular orbit and firing the second stage engine twice. The first stage booster failed to land on the drone ship[504] due to incorrect wind data.[505] This was the first time a flight proven booster failed to land.								
82	7 March 2020, 04:50[506]	F9 B5 △ B1059.2	CCAFS, SLC-40	SpaceX CRS-20 (Dragon C112.3 △)	1,977 kg (4,359 lb)[507]	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
	Last launch of phase 1 of the CRS contract. Carries *Bartolomeo*, an ESA platform for hosting external payloads onto ISS.[508] Originally scheduled to launch on 2 March 2020, the launch date was pushed back due to a second stage engine failure. SpaceX decided to swap out the second stage instead of replacing the faulty part.[509] It was SpaceX's 50th successful landing of a first stage booster, the third flight of the Dragon C112 and the last launch of the cargo Dragon spacecraft.								
83	18 March 2020, 12:16[510]	F9 B5 △ B1048.5	KSC, LC-39A	Starlink 5 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Failure (drone ship)
	Fifth operational launch of Starlink satellites. It was the first time a first stage booster flew for a fifth time and the second time the fairings were reused (Starlink flight in May 2019).[511] Towards the end of the first stage burn, the booster suffered premature shut down of an engine, the first of a Merlin 1D variant and first since the CRS-1 mission in October 2012. However, the payload still reached the targeted orbit.[512] This was the second Starlink launch booster landing failure in a row, later revealed to be caused by residual cleaning fluid trapped inside a sensor.[513]								
84	22 April 2020, 19:30[514]	F9 B5 △ B1051.4	KSC, LC-39A	Starlink 6 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)

Figure 1: Web_Scraping

Figure 1: HTML Tables showing list of Falcon 9 and Falcon Heavy launches.

```
df= pd.DataFrame({ key:pd.Series(value) for key, value in launch_dict.items() })
df.head()
```

	Flight No.	Launch site	Payload	Payload mass	Orbit	Customer	Launch outcome	Version Booster	Booster landing	Date	Time
0	1	NaN	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success\n	F9 v1.0B0003.1	Failure	4 June 2010	18:45
1	2	NaN	Dragon	0	LEO	NASA	Success	F9 v1.0B0004.1	Failure	8 December 2010	15:43
2	3	NaN	Dragon	525 kg	LEO	NASA	Success	F9 v1.0B0005.1	No attempt\n	22 May 2012	07:44
3	4	NaN	SpaceX CRS-1	4,700 kg	LEO	NASA	Success\n	F9 v1.0B0006.1	No attempt	8 October 2012	00:35
4	5	NaN	SpaceX CRS-2	4,877 kg	LEO	NASA	Success\n	F9 v1.0B0007.1	No attempt\n	1 March 2013	15:10

Figure 2: Web_Scraping

Figure 2: Result Pandas DataFrame from BeautifulSoup webscraping.

Course 5: Python Project for Data Science

Overview: Work on a real-world data science project using Python. **Key Project:** Apply Python fundamentals and data structures to analyze data.

Course 6: Databases and SQL for Data Science with Python

Overview: Analyze and manage data using SQL and Python. **Key Project:** Create and query relational databases, utilizing advanced SQL techniques.

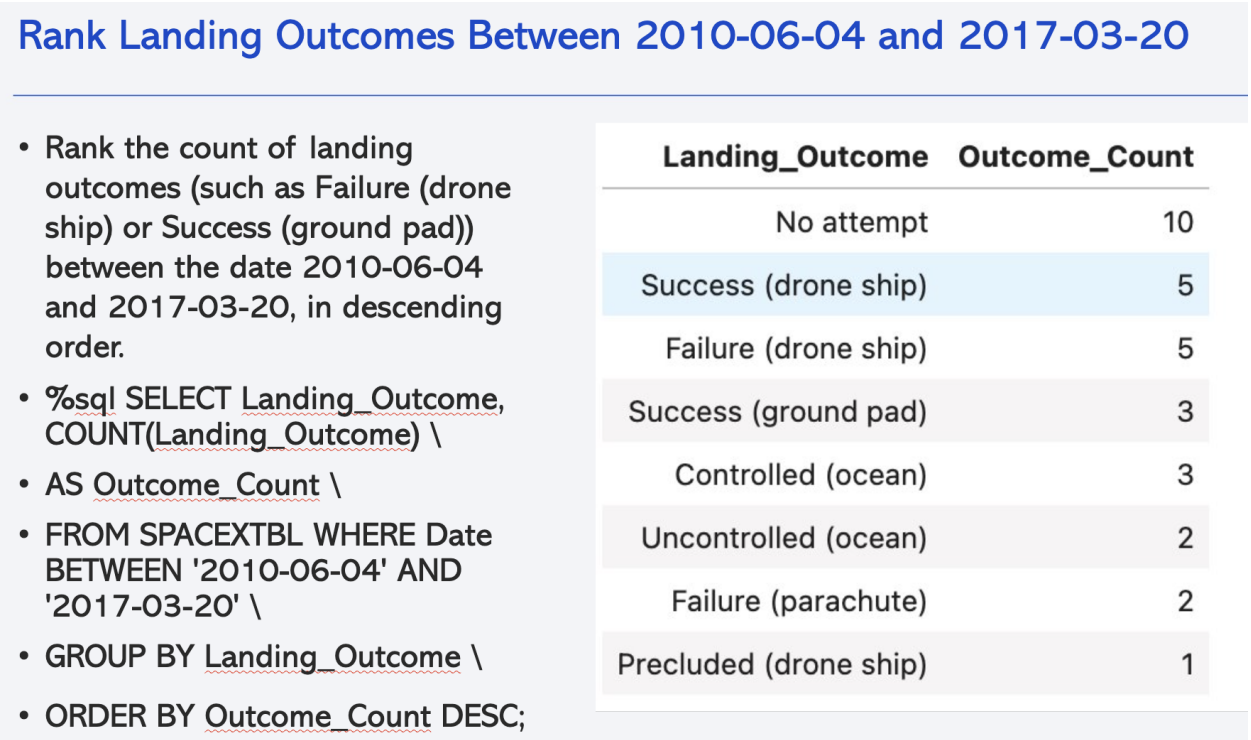


Figure 3: sql

Figure 3: SQL analysis ranking landing outcomes between 2010-06-04 and 2017-03-20.

Course 7: Data Analysis with Python

Overview: Perform data cleaning, preparation, and exploratory analysis. **Key Project:** Manipulate and analyze datasets using Pandas, Numpy, and Scipy.

Course 8: Data Visualization with Python

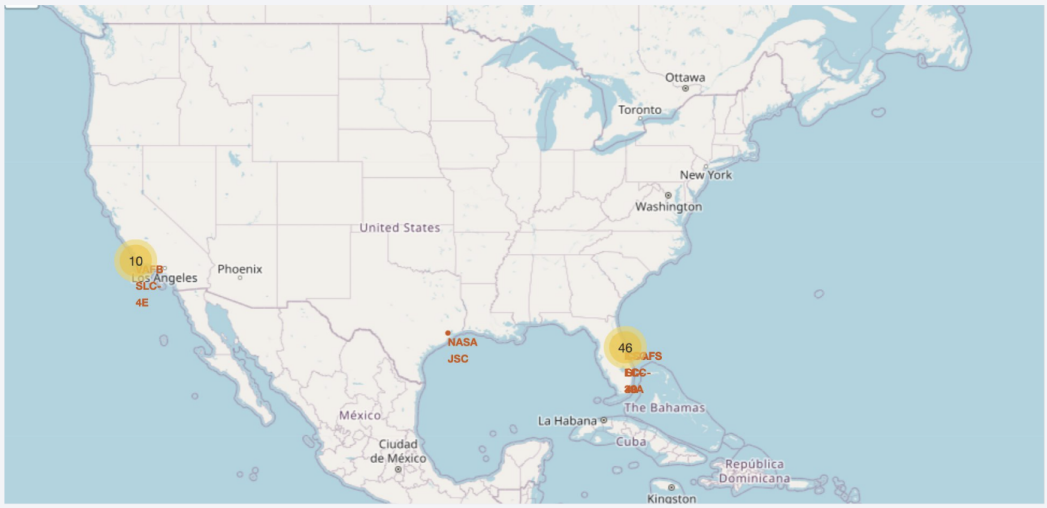
Overview: Implement data visualization techniques to tell compelling stories. **Key Project:** Create various charts and plots using Matplotlib, Seaborn, and Folium.

Figure 4: Interactive folium map showing SpaceX launch sites.

Figure 4: Interactive folium map showing launch success/failures at a specific launch site using marker-clusters.

Figure 5: Plotly interactive dashborad showing launch successes by launch site. Kennedy Space Center has the highest successful launch rate.

Interactive Map of SpaceX Launch Sites



- SpaceX launch sites. Note: i) proximity to equator (capitalize on Earth's rotational orbit for reduced fuel cost to achieve orbit) and ii) proximity to coast (reduced risk of injury to person or property).

Figure 4: sql

Interactive Map Showing Successes/Failures

- Kennedy Space Center successes (Green) and failures (Red) shown on interactive folium map.

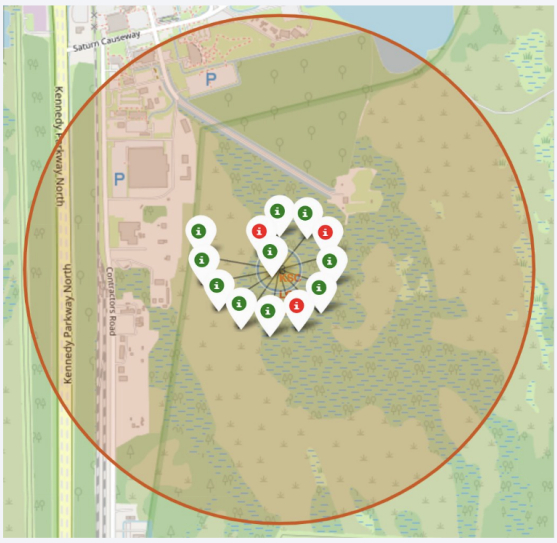


Figure 5: SQL

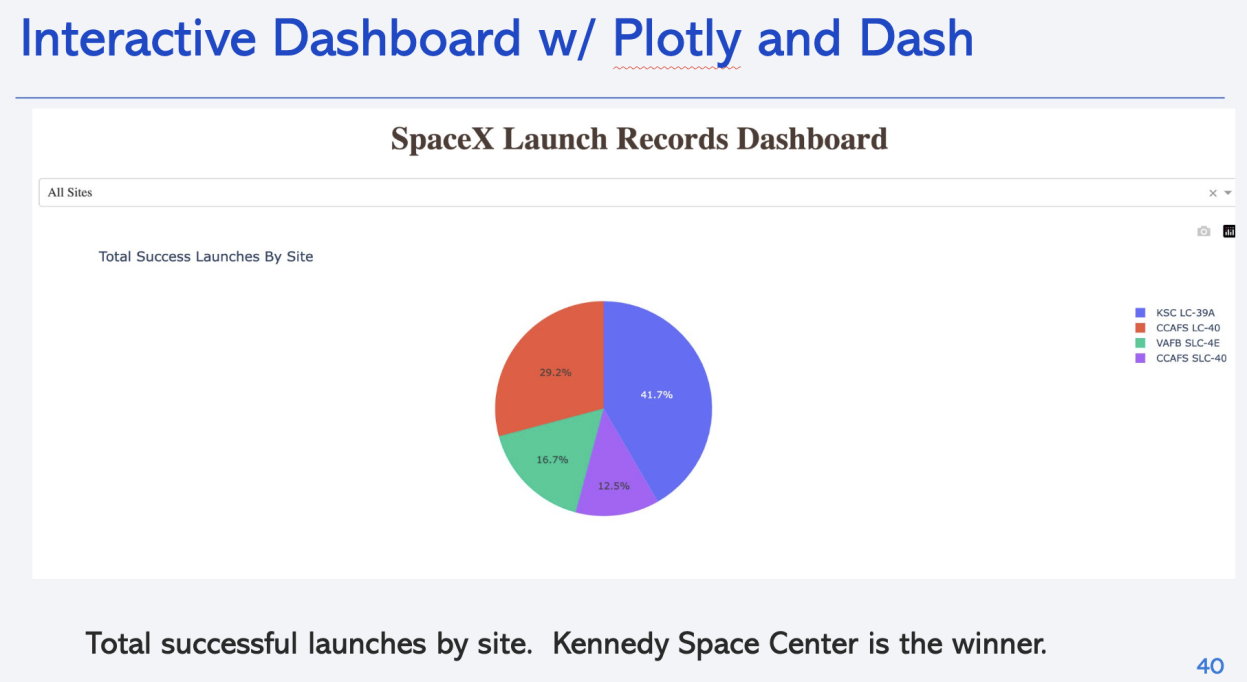


Figure 6: Dashboard

Course 9: Machine Learning with Python

Overview: Understand and implement machine learning algorithms. **Key Project:** Write Python code for classification techniques like KNN and decision trees.

Figure 6: Classification model accuracy comparison: Decision Trees is the winner.

Course 10: Applied Data Science Capstone

Overview: Demonstrate proficiency through a comprehensive capstone project. **Key Project:** Perform data wrangling, EDA, model development, and evaluation.

Course 11: Generative AI: Elevate Your Data Science Career

Overview: Leverage generative AI tools to enhance data science workflows. **Key Project:** Generate and augment datasets using tools like GPT-3.5 and tomat.ai.

Course 12: Data Scientist Career Guide and Interview Preparation

Overview: Prepare for a data science career with job search strategies and interview tips. **Key Project:** Build a resume, portfolio, and practice interview techniques.

Conclusion

This portfolio not only demonstrates my technical skills and knowledge but also my ability to apply them to real-world problems. Each project is a testament to my journey through the IBM Data Science Professional Certificate program and my readiness to excel as a data scientist.

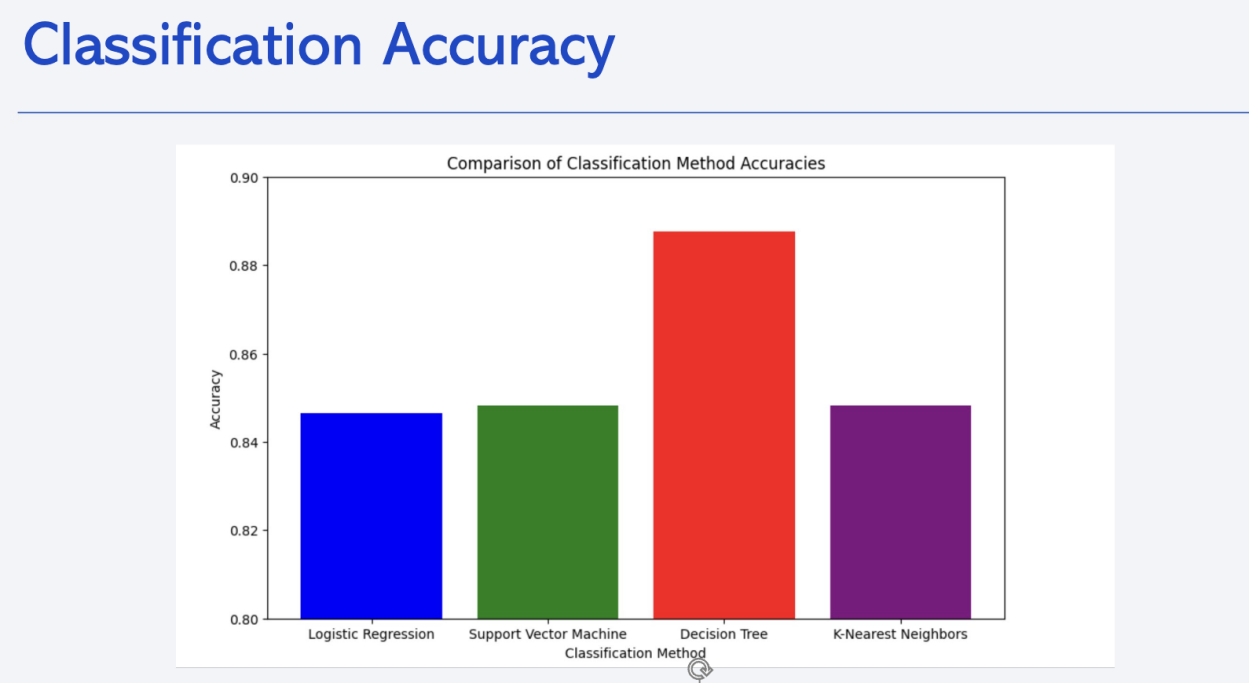


Figure 7: Classification Model Comparison

Explore the projects and witness the impact of data science in action!

Certificate Verification

Certificate #: Q5Q6H87SM9S9