**Applied Data Science Capstone**

Table of Contents

[**Week 1: Introduction** 3](#_Toc128174214)

[*Welcome* 3](#_Toc128174215)

[*Course Introduction* 3](#_Toc128174216)

[*Capstone Introduction and Understanding the Datasets* 3](#_Toc128174217)

[*Project Scenario and Overview* 3](#_Toc128174218)

[*Obtain IBM Cloud Feature Code and Activate Trial Account* 3](#_Toc128174219)

[*IBM Watson Account Creation and Watson Studio* 3](#_Toc128174220)

[*Getting started with GitHub* 3](#_Toc128174221)

[*Publishing Notebook to Github* 3](#_Toc128174222)

[*IBM Watson Setup and Project Creation* 3](#_Toc128174223)

[*Adding a Notebook to the Project* 3](#_Toc128174224)

[*Collecting the Data* 4](#_Toc128174225)

[*Data Collection Overview* 4](#_Toc128174226)

[*Hands-on Lab: Complete the Data Collection API Lab* 4](#_Toc128174227)

[*[Optional] Hands-on Lab: Complete the Data Collection API Lab* 4](#_Toc128174228)

[*Data Collection with Web Scraping* 4](#_Toc128174229)

[*[Optional] Hands-on Lab: Complete the Data Collection with Web Scraping Lab* 4](#_Toc128174230)

[*Check Points: Data Collection API* 4](#_Toc128174231)

[*Graded Quiz: Data Collection API with Web Scraping* 4](#_Toc128174232)

[*Data Wrangling* 4](#_Toc128174233)

[*Data Wrangling Overview* 4](#_Toc128174234)

[*Hands-on Lab: Data Wrangling Overview* 5](#_Toc128174235)

[*[Optional] Hands-on Lab: Data Wrangling Overview* 5](#_Toc128174236)

[*Check Points: Data Wrangling* 5](#_Toc128174237)

[*Graded Quiz: Data Wrangling Quiz* 5](#_Toc128174238)

[**Week 2: Exploratory Data Analysis (EDA)** 5](#_Toc128174239)

[*Exploratory Analysis Using SQL* 5](#_Toc128174240)

[*Course Introduction* 5](#_Toc128174241)

[*Project Scenario and Overview* 5](#_Toc128174242)

# **Week 1: Introduction**

## *Welcome*

### *Course Introduction*

Hello! And welcome to this capstone course.

Congratulations for making it this far! My name is Joseph Santarcangelo, Yan Luo and Azim Hirjani . We are pleased to be your instructors and course developers for this capstone course. You will apply your data science skills as a Data scientist for a private space launch company in this project.

As a starting point of almost all data science projects, you need to collect data, as much and relevant as possible.

You will be collecting data from various sources. After your raw data has been collected, you will need to improve the quality by performing data wrangling.

Then you can start exploring the processed data. We will be your guide as we explore some really interesting real-world datasets together. You'll get to practice your SQL skills as we query the data and gather insights.

You'll gain further insights into the data by applying some basic statistical analysis and data visualization, you'll be able to see directly how variables might be related to each other.

We'll drill down into finer levels of detail by splitting the data into groups defined by categorical variables or factors in your data.

You will be guided to build, evaluate, and refine predictive models for discovering more exciting insights.

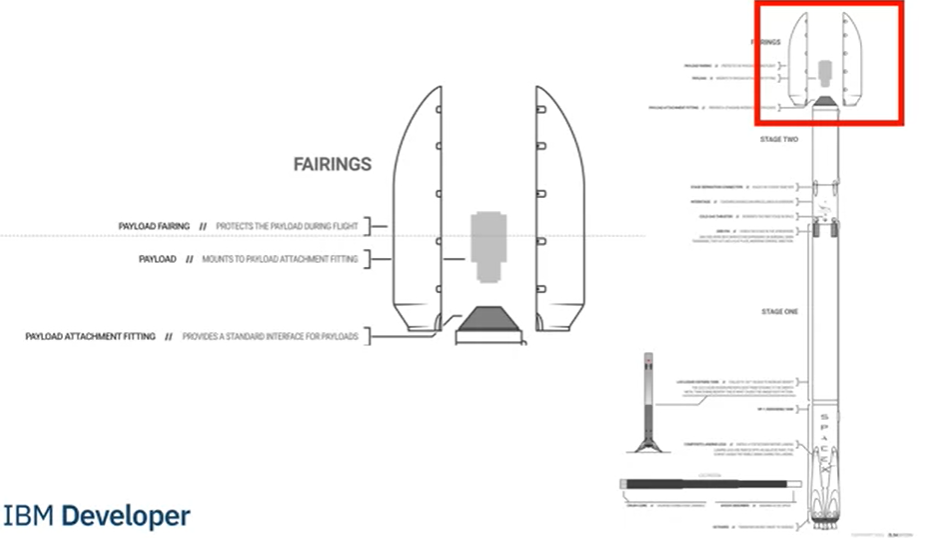
The final task of this capstone project is to create a presentation that will be developed into stories of all your analysis.

Thanks and good luck!

## *Capstone Introduction and Understanding the Datasets*

### *Project Scenario and Overview*

In this video, we will provide the Project Scenario and Overview. The commercial space age is here, companies are making space travel affordable for everyone. Virgin Galactic is providing suborbital spaceflights. Rocket Lab is a small satellite provider. Blue Origin manufactures sub-orbital and orbital reusable rockets. Perhaps the most successful is SpaceX. SpaceX’s accomplishments include: Sending spacecraft to the International Space Station. Starlink, a satellite internet constellation providing satellite Internet access. Sending manned missions to Space. One reason SpaceX can do this is the rocket launches are relatively inexpensive. SpaceX advertises Falcon 9 rocket launches on its website with a cost of 62 million dollars; other providers cost upwards of 165 million dollars each, much of the savings is because SpaceX can reuse the first stage. Therefore, if we can determine if the first stage will land, we can determine the cost of a launch. Spaces X’s Falcon 9 launch like regular rockets. To help us understand the scale of the Falcon 9, we are going to use these diagrams from Forest Katsch, at  zlsadesign.com. He is a 3D artist and software engineer. He makes infographics on spaceflight and spacecraft art. He also makes software. The payload is enclosed in the fairings.



Stage two, or the second stage, helps bring the payload to orbit,

Diagram

Description automatically generated

but most of the work is done by the first stage. The first stage is shown here. This stage does most of the work and is much larger than the second stage. Here we see the first stage next to a person and several other landmarks. This stage is quite large and expensive.

Diagram

Description automatically generated

Unlike other rocket providers, SpaceX's Falcon 9 Can recover the first stage. Sometimes the first stage does not land. Sometimes it will crash as shown in this clip. Other times, Space X will sacrifice the first stage due to the mission parameters like payload, orbit, and customer. In this capstone, you will take the role of a data scientist working for a new rocket company. Space Y that would like to compete with SpaceX founded by Billionaire industrialist Allon Musk. Your job is to determine the price of each launch. You will do this by gathering information about Space X and creating dashboards for your team. You will also determine if SpaceX will reuse the first stage. Instead of using rocket science to determine if the first stage will land successfully, you will train a machine learning model and use public information to predict if SpaceX will reuse the first stage.

### *Obtain IBM Cloud Feature Code and Activate Trial Account*

### *IBM Watson Account Creation and Watson Studio*

### *Getting started with GitHub*

### *Publishing Notebook to Github*

### *IBM Watson Setup and Project Creation*

### *Adding a Notebook to the Project*

## *Collecting the Data*

### *Data Collection Overview*

### *Hands-on Lab: Complete the Data Collection API Lab*

### *[Optional] Hands-on Lab: Complete the Data Collection API Lab*

### *Data Collection with Web Scraping*

### *[Optional] Hands-on Lab: Complete the Data Collection with Web Scraping Lab*

### *Check Points: Data Collection API*

### *Graded Quiz: Data Collection API with Web Scraping*

## *Data Wrangling*

### *Data Wrangling Overview*

### *Hands-on Lab: Data Wrangling Overview*

### *[Optional] Hands-on Lab: Data Wrangling Overview*

### *Check Points: Data Wrangling*

### *Graded Quiz: Data Wrangling Quiz*

# **Week 2: Exploratory Data Analysis (EDA)**

## *Exploratory Analysis Using SQL*

### *Course Introduction*

### *Project Scenario and Overview*