



**TASK**

# **Capstone Project - Data Analysis**

Visit our website

# Introduction

## WELCOME TO THE CAPSTONE PROJECT - DATA ANALYSIS TASK!

This capstone project will consolidate what you have learnt thus far and allow you to put your newly acquired knowledge concerning exploratory analysis to the test.

## Compulsory Task 1

Create a copy of the **Capstone\_Project\_template.ipynb** file and save it as **movies.ipynb**.

Follow the instructions to clean, sanitise, and explore the **movies.csv** dataset.

Compile an EDA on the movies dataset in the **Movies EDA Task Doc** explaining your visualisations, investigations, and findings.

To do this, you will need to:

- Load the dataframe in
- Clean the data
- Remove duplicate rows
- Discard entries with a zero movie budget
- Manipulate certain columns to the correct data type
- Answer the questions about the data

## Compulsory Task 2

Create another copy of the **Capstone\_Project\_template.ipynb** file and save it as **automobile.ipynb**.

Recreate what you did in Compulsory Task 1 to clean, sanitise, and explore the **automobile.txt** dataset.

Adapt the narrative in the text cells to be relevant to the automobile dataset. Make sure to update the explanations and descriptions to accurately reflect the data and analysis being performed.

Compile an EDA on the automobile dataset in the **Automobile EDA Task Doc** explaining your visualisations, investigations, and findings.

Follow the same steps as you did for the previous task to visualise and explore the dataset thoroughly.

## Completed the task(s)?

Ask an expert to review your work!

[Review work](#)



Rate us

**Share your thoughts**

HyperionDev strives to provide internationally-excellent course content that helps you achieve your learning outcomes.

Think that the content of this task, or this course as a whole, can be improved? Do you think we've done a good job?

[Click here](#) to share your thoughts anonymously.

---

## REFERENCES

Bourke, D. A Gentle Introduction to Exploratory Data Analysis. Retrieved January 13, 2019, from towardsdatascience.com:

<https://towardsdatascience.com/a-gentle-introduction-to-exploratory-data-analysis-f1d843b8184>

NIST. (n.d.). What is EDA? Retrieved May 6, 2019, from Engineering Statistics Handbook:

<https://www.itl.nist.gov/div898/handbook/eda/section1/eda11.htm>

Wikipedia. Exploratory data analysis. Retrieved April 27, 2019, from Wikipedia:

[https://en.wikipedia.org/wiki/Exploratory\\_data\\_analysis](https://en.wikipedia.org/wiki/Exploratory_data_analysis)