

**Title: Interactive Charts with Altair Using IMDB Movies Data sets**

Faculty of Computer Science

Dublin business School

Group Assignments2 on Visualization

|  |  |  |  |
| --- | --- | --- | --- |
| **Author:**  **Student No:** | **Opeyemi Salami**  **10387356** | Supervisor: | Basel Magableh |
| E-mail: | 10387356@mydbs.ie | E-mail: | Basel.magableh@dbs.ie |
| **Author:**  **Student No:**  **Email:** | **Oluwatoyin Kolade**  **10392388** |  |  |
| **Author:**  **Student No:**  **Email** | **Ekpemhi Ojeifo**  **10393484** | : |  |

**IMDB Movies Data sets**

Altair is a declarative statistical visualization library for Python, based on [Vega](http://vega.github.io/vega) and [Vega-Lite](http://vega.github.io/vega-lite). Altair offers a powerful and concise visualization grammar that enables you to build a wide range of statistical visualizations quickly. It is Altair is designed to work with data in the form of Panda data frames, and contains a loader for this and other built-in datasets:

The aim of this project is to design an interactive graphics to communicate a

Compelling insight into the datasets that were used by performing exploratory

analysis on the datasets. Techniques such as panning, zooming, brushing,

details-on-demand (tooltips) are used to draw focus to an item. In this project we will create visualizations to explore Altair in-built data namely – “Seattle weather” and “Movies”.

This project was splinted into three different categories which was completed among the team members - The Interactive charts using Altair, Altair chart configuration and Altair transformation. It took about three weeks for us to complete the project because of the challenges we had along the line. As a team, we concluded on plotting interactive charts with possible techniques such as panning, zooming, brushing, details-on-demand (e.g., tooltips). Throughout this series, we’ll be working with the Seattle weather and movies datasets

IMDB movies is the world most and popular and authoritative source for movie, TV and celebrity content. IMDb offers a rating scale that allows users to rate films on a scale of one to ten User ratings of films. As one adjunct to data, IMDb indicates that submitted ratings are filtered and weighted in various ways in order to produce a weighted mean that is displayed for each film, series.

**Insights:**

Predicting Ratings on number of votes

Movies success based on a certain metric

What movies tend to get higher vote counts and vote averages on IMDB?

Creating an interactive exploratory chart to visualize the highest IMDB ratings,

Content Based charts to view most focused Creative \_Type.

Seattle weather dataset is an in-built data in Altair with columns for the temperature (in Celsius), precipitation (in millimeters), wind speed (in meter/second), and weather type. We have one row for each day from January 1st, 2012 to December 31st, 2015