

Group Project

Team: AWALONG

October 26, 2017

1 Project Title

The studies of the influence factors of movie box office

2 Project Description

Our team is going to study the influence factors of movie box office by using various datasets. like the reviews of a movie, the actors in a movie, the type of the movie and so on.

Data Collection: Data source: Box office, IMDB, Twitter and etc.

Techniques: python scrapy, BeautifulSoup modules.

Data Cleaning: Techniques: python numpy, pandas modules.

Data Analysis: Techniques: python numpy, pandas, scikit-learn.

Data visualization: Techniques python matplotlib, seaborn.

3 Team Structure

Team Leader:

Di Wu: dw4y17@soton.ac.uk

Deputy:

Han Tong: htzy17@soton.ac.uk

Other members:

Zelong Yan: zy4y17@soton.ac.uk Yi Shen: ys4a17@soton.ac.uk

Qiling Chen: ql1y17@soton.ac.uk

4 Git Repository Address

<https://github.com/awalongsoton/Data-Science>