



BARCELONA SCHOOL OF ECONOMICS

DATA SCIENCE METHODOLOGY PROGRAM

22DM014 - INTRO TO NATURAL LANGUAGE PROCESSING

Analysis of Amazon Books and Reviews

Authors:

CHEN, Joshua
EL DAOU, Mahmoud
GALLEGOS, Rafael

Professors:

MUELLER, Hannes
GARCIA, Luis

Contents

1	Introduction	2
2	Literature Review	2
3	Dataset	2
4	Preprocessing	2
5	Exploratory Data Analysis	2
6	Predicting Review Score from Text	2
7	Review Description Alignment	2
8	Conclusions	2

List of Figures

List of Tables

- 1 Introduction
- 2 Literature Reviews
- 3 Dataset
- 4 Preprocessing
- 5 Exploratory Data Analysis
- 6 Predicting Review Score from Text
- 7 Review Description Alignment
- 8 Conclusions