Faculté des Sciences appliquées

MATH2021-1 High-dimensional statistics

Project 1 : Exploratory Data Analysis

Teacher :
Gentiane HAESBROECK

Group :
Romain LAMBERMONT
Arthur LOUIS

October 26, 2022

Contents

1 Presentation of the data **1**
1.1 Discussion on the data . . . 1
1.2 Link between the variables . . . 1

2 Information about missing data **1**

3 Exploratory data analysis **1**
3.1 Statistical analysis . . . 1
3.2 Graphical analysis . . . 1
3.3 Correlation structure of the data . . . 1
3.4 Outlying observations using Mahalanobis distance . . . 1

4 Correlation analysis with data reduction **1**
4.1 Choice between PCA and t-SNE . . . 1
4.2 2D plot of the data . . . 1

List of Figures

List of Tables

1 Presentation of the data

1.1 Discussion on the data

1.2 Link between the variables

2 Information about missing data

3 Exploratory data analysis

3.1 Statistical analysis

3.2 Graphical analysis

3.3 Correlation structure of the data

3.4 Outlying observations using Mahalanobis distance

4 Correlation analysis with data reduction

4.1 Choice between PCA and t-SNE

4.2 2D plot of the data