Ruprecht-Karls-Universität Heidelberg
Fakultät für Biowissenschaften
Bachelorstudiengang Molekulare Biotechnologie

ksdflsdjf sfdsfd sfsf

Data Science Project SoSe 2022

Autoreb	Max Mustermann, jklsdfksldkfslkdjf
Geburtsort	sdfjksafdls
Abgabetermin	20.07.2022

Contents

1 Introduction

1.1 Computational Tools

1.1.1 Dimension reduction

1.1.1.1 PCA

1.1.1.2 UMAP

1.1.2 Statistical analysis

1.1.2.1 Shapiro-Wilks test

Shapiro-Wilks test is a normality test based on regression and correlation. It tests the null hypothesis that the data follows a normal distribution. Small values of SW test statistic indicate no normality of the data thus the null hypothesis is rejected. SW values of one suggest normality Yap and Sim (2011).

1.1.2.2 Wilcoxon rank-sum and signed-rank test

Wilcoxon rank-sum test and Wilcoxon signed-rank test both are non-parametric statistical hypothesis tests that can be used when the data does not follow a normal distribution. Wilcoxon signed-rank test is used to analyze matched-pair or one-sample data. It tests the null hypothesis that there is no difference in probability distribution of first and second sample, hence the distribution of pairwise differences is centered at zero. The test is based on ranked absolute values of differences Woolson (2007). Wilcoxon rank-sum test is performed when analyzing unpaired-data and is likewise based on ranked values. The null hypothesis states that there is no association between the two samples Rey and Neuhäuser (2011).

1.1.2.3 H-test

1.1.3 Clustering

1.1.3.1 Kmeans

1.1.3.2 Hierarchial clustering

1.1.4 GSEA

1.1.5 Regression

2 Materials and Methods

2.1 Data cleaning

2.2 TCGA pan-cancer analysis

2.3 KIRC specific analysis

2.4 Packages

3 Results

3.1 TCGA pan-cancer analysis

3.2 KIRC specific analysis

4 Discussion

5 Concluding remarks/Outlook

6 References

Rey, D, and Neuhäuser, M (2011). Wilcoxon-signed-rank test. In: International Encyclopedia of Statistical Science, ed. M Lovric, Berlin, Heidelberg: Springer Berlin Heidelberg, 1658–1659.

Woolson, RF (2007). Wilcoxon signed-rank test. Wiley Encyclopedia of Clinical Trials, 1–3.

Yap, BW, and Sim, CH (2011). Comparisons of various types of normality tests. Journal of Statistical Computation and Simulation 81, 2141–2155.

7 Appendix