

# Introduction

*Frederikke Haider*

*20 august 2016*

## Introduction

As most students, both former and present might recognize as one of the big questions while studying is; “ *Which subject should I choose and how will I do on my exam?*”. In the ideal world, the result on one’s exam should only depend on ability, interest, and how well prepared you are, but is this always the case? “*Could the type of exam assessment explain whether you get a good grade or not?*”

This is the question we will try to answer in this paper by taking advantage of the development in the field of Social Science Data, which has opened up many new possibilities for gathering, manipulate and investigate data using the internet.

Our data is gathered from the University of Copenhagen’s web page (<http://karakterstatistik.stads.ku.dk>) that shows the grade distribution for the different subjects across all faculties and institutes dating back to summer 2011. There are many opportunities and different approaches that could be used analysing the available data and to try and answer our question. In this paper however, the main focus will be on the department of Economics, and if the grade distribution differs depending on the type of exam (written, oral or take home) and/or if you attend the ordinary or re-exam.

You will be able to see how large amounts of information can be gathered from the internet, and how different techniques such as merging, analysing and visualizing helped us answer our question. We will go through the process of scraping, cleaning and modelling, and look at the different problems we faced along the way and how we solved them. Finally we will discuss different approaches that could be used and conclude on our question.