

Exam project description – Group 20

Online costumer reviews are important tools when navigating as a tourist in a foreign city.

An important element of any holiday is food and drinks! This is where an app/site like TripAdvisor comes in play.

We will scrape data from the travel portal TripAdvisor's homepage, on restaurants in the Copenhagen area. We will focus on if the distance to the city center of Copenhagen affects the quality of the restaurants, if the level of service is affected by the distance to the city center and if there is a direct link between price-level and costumer ratings. We will make use of spacious visualizations (maps), linear models and we will investigate using machine learning tools if it is possible to predict the ratings based on the characteristics of the restaurants.

Methods:

1. We will create a tool that scrapes TripAdvisor site for restaurants in the Copenhagen area. Next it will give us a list of unique links to the individual restaurants. Last we will scrap each restaurant for the data we need.
2. Data manipulation after the scraping is finished.
3. Visualize a map using the geo data scraped.