Introduction

This report depicts a research performed regarding active Danish real-estate property valuations from 2016-2019. Since 2011, the yearly number of real estate properties sold has slightly increased, with 2017 being the year with most sold properties in 10 years (Danmarks Statistik, 2018). The price of real estate properties is also rising through 2019 (Danmarks Statistik, 2019). Although this generally depicts a willingness of buyers to pay more for real-estate property, and higher valuations from agents, a closer analysis of current and active valuations contributes with local view of which conditions contribute to the valuation of a real-estate property.

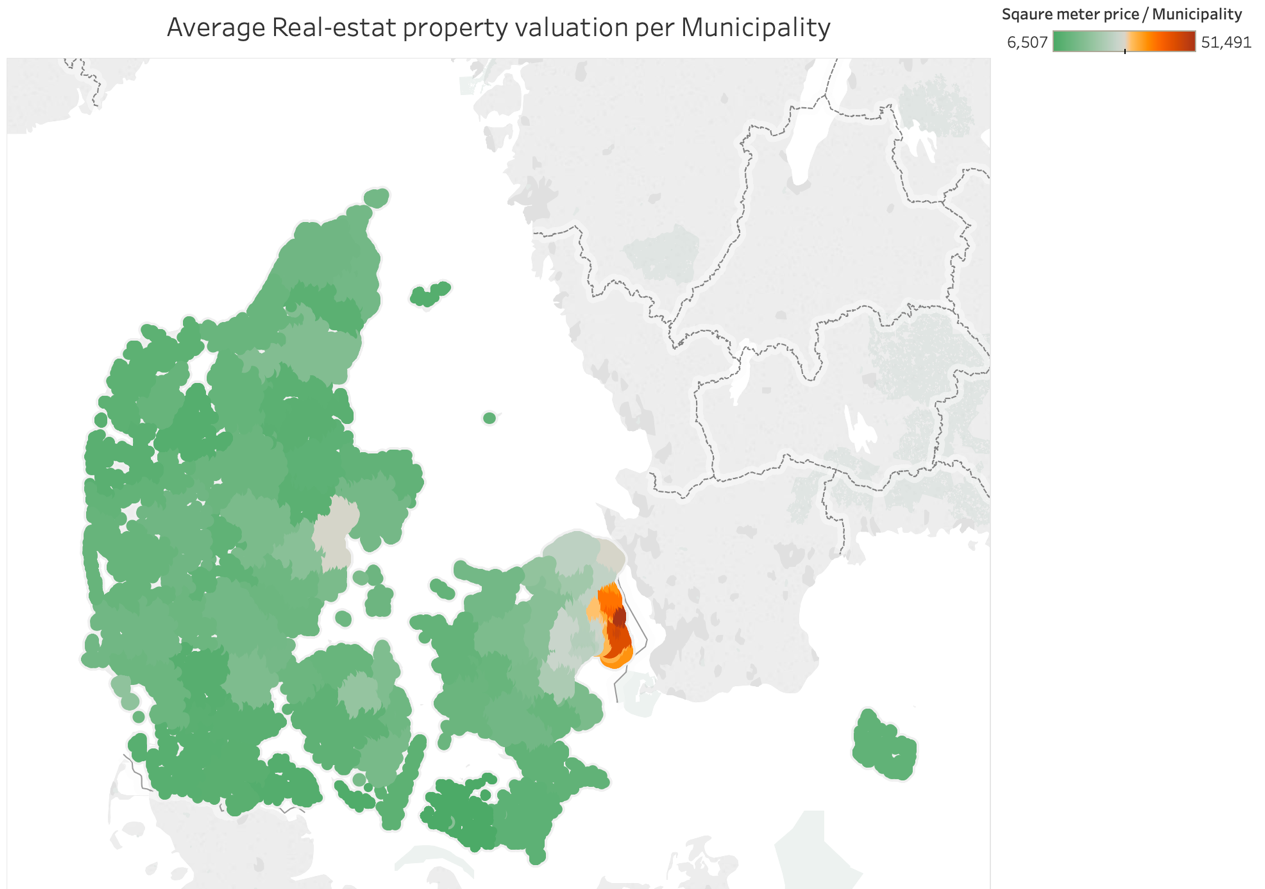


Figure XX

Figure XX shows the average square meter price valuation of active offers per municipality in Denmark. The maximum square meter price being approximately 8 times higher in the municipality of Gentofte, than Lolland, the municipality with lowest average square meter price valuations. A tendency visualized in the figure, is that average square meter price valuations are higher in in highly populated municipalities and in suburban areas surrounding Copenhagen.

The valuations included in the research are collected from one of Denmarks largest online real-estate websites named Boliga.dk. The Boliga data contains approximately 66 thousand active offerings with valuations ranging from 15 thousand to 85 million. This research paper intends to analyze real estate price valuations and which effects geo- and sociodemographic criteria affect valuations, using machine learning models. Our research questions is as follows:

*Which features are most relevant for predicting evaluations of real estate properties in Denmark?*

The research paper contains a section describing the construction of research data was performed explains the choices of gathering meaningful features for the machine learning model. Also, a section revolving the choice and optimization of machine learning models is included, where the intention is to provide insights on our progress on finding the optimal model. As a result, a preferable machine learning model is chosen with a discussion of its usability.