Continuous Assessment 2

Author: Ciara O'Hara

Programme: MSc in Data Analytics May 2023

Student ID: sbs23015

Email: sbs23015@student.cct.ie

May 6, 2023

Contents

0.1 brief

You have been tasked with analysing Ireland's Construction data and comparing the Irish Construction sector with other countries worldwide. This analysis should also include forecasting, sentiment analysis and evidence-based recommendations for the sector as well as a complete rationale of the entire process used to discover your findings. Your Research could include export, import, trade imbalance, house production, material stock, labour/skill pool, etc. (or any other relevant topic EXCEPT Climate change) with Ireland as your base line.

Abstract

Introduction

Info:

https://oecdecoscope.blog/2021/12/13/finlands-zero-homeless-strategy-lessons-from-a-success-story/comment-page-1/

https://www.oecd.org/housing/policy-toolkit/data-dashboard/boosting-efficiency/

https://www.linesight.com/insights/regional-report/europe-2021/

https://pxdata.stat.fi/PxWeb/pxweb/en/StatFin/

How to extract the API data: $https://www.scb.se/api_en/$

$https://github.com/kirajcg/pyscbwrapper/blob/master/pyscbwrapper_en.ipynb$

********* Twitter ********* HousingForAll (Ireland) homeless

Materials and Methods

First steps were to determine and identify an appropriate Irish dataset under the theme of Constrction. There were two datasets found: A House Construction Cost Index from 1975 - 2017, and a social housing construction status reports from 2017 - 2021. Homelessness in Ireland is an issue of major significance and public importance at the moment. In fact, this is an issue across Europe to varying degrees, with the excpetion of Finland. It was decided to try to determine - from publicly available data - what Finland has done differently, the factors that may have impacted that, and to attempt a sentiment analysis around the topic in both Ireland and Finland. As such, the next step was to gather appropriate and complementary Finnish data. Statistics Finland's free-of-charge statistical databases, Tilastokeskus was found, which included ... in JSON format, accessed via APIs.

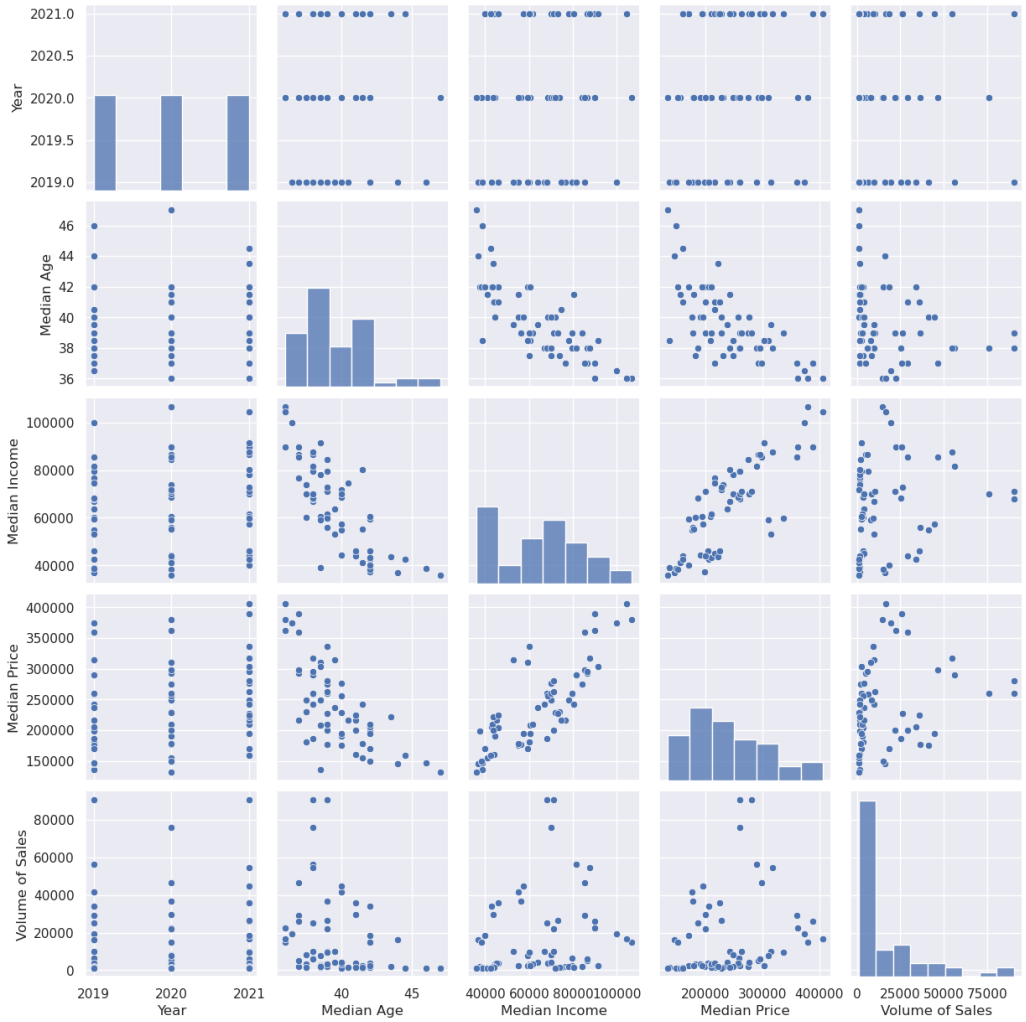

Figure 3.1: *Sample figure.*

3.1 Data sources

3.2 Programming

3.2.1 Python libraries used

3.3 Data preparation

3.3.1 Cleaning the data

3.3.2 Data visualisation

3.4 Statistics

3.4.1 Python libraries

3.4.2 Statistical analysis

Table 3.1: *Table template.*

Variable	W Test statistic	p-value
Median Age	0.93	3.18×10^{-35}
Median Income	0.94	2.22×10^{-33}
Median Price	0.87	4.34×10^{-44}
Volume of Sales	0.86	0

3.5 Machine learning

3.5.1 Machine learning methods used

Results

Conclusion