**Data Journal – UK Housing EDA**

|  |  |
| --- | --- |
| **Date:** 25/1/24 | **Course/topic: Picking a topic** |
| **Prompt:** | * UK housing exploratory data analysis * News articles * Datasets |
| **Journal Entry:** | I hope to use this project to build my confidence in using SQL, python, pandas, seaborn, and Tableau. I might encounter issues in all these areas as I’m still a beginner. But I hope to be able to demonstrate good knowledge of the data analysis fundamentals to make the first project for my online portfolio on GitHub. I have collected news articles to help frame the problem my project will address. I have also collected some datasets that I believe will be useful, some Python tutorials and a data journalism article example:   * https://www.newyorker.com/tech/annals-of-technology/mapping-new-york-noise-complaints (example of data journalism) * https://www.theguardian.com/business/2024/jan/05/not-that-difficult-buy-home-uk-natwest-chair-howard-davies (news article) * https://www.kaggle.com/code/kanncaa1/data-sciencetutorial-for-beginners/notebook (Python tutorial) * https://www.kaggle.com/code/benhamner/python-data-visualizations/notebook (Python tutorial) * https://www.kaggle.com/code/mammadabbasli/bank-marketing-campaign/input (notebook example) * https://datasetsearch.research.google.com/search?src=0&query=housing%20uk&docid=L2cvMTFxNDFndnYxYw%3D%3D (datasets) * https://www.kaggle.com/datasets/liamhealy/uk-mortgage-rates (dataset) * https://www.ft.com/content/f21642d8-da2d-4e75-886e-2b7c1645f063 (news article) * https://www.theguardian.com/society/2023/oct/29/right-to-buy-housing-crisis-home-ownership-britain-property-rowan-moore (news article) * https://www.telegraph.co.uk/business/2024/01/25/eco-fanatics-are-creating-a-new-housing-crisis/ (news article) * https://www.theguardian.com/commentisfree/2023/jul/05/uk-housing-crisis-mortgages-private-local-allowance-freeze-tenants (news article) * https://www.ft.com/content/c7abc82c-8d1e-4ec9-8630-971fd806cc54 (news article) * https://www.forbes.com/uk/advisor/mortgages/2024/01/30/mortgage-updates/ (news article) |
| **Other thoughts or questions:** | Use Jupyter notebooks and python for bulk of project (including pandas, matplotlib, and seaborn). Explore some SQL at beginning. Explore some Tableau at end. Submission at UofG requires a data-journalism article and final report – 2 word documents. It will upload my notebook to GitHub. New articles show housing is an important contemporary topic to discuss and appeals to a large audience. Datasets show there is a large amount of availability on my topic, so I won’t need to perform any of my own data collection. These are trusted open data providers (the government). |

|  |  |
| --- | --- |
| **Date:** 30/1/24 | **Course/topic: Writing SOW** |
| **Prompt:** | Scope of Work |
| **Journal Entry:** | I wrote my scope of work document to flesh out my ideas for the project in more detail. It covered what my project would discuss, and what it would not. I made a timeline and major milestones. I also included what the deliverables would be. |
| **Other thoughts or questions:** | Still not sure exactly what data I will analyse. Housing prices and mortgages yes. But I need to learn some financial knowledge and housing market knowledge to gain a better understanding of the context of the data. |

|  |  |
| --- | --- |
| **Date:** 31/1/24 | **Course/topic: Starting Notebook** |
| **Prompt:** | Creating Notebook  Housing price dataset  Visualisation |
| **Journal Entry:** | I created the notebook for my project. I included the tech stack I plan to use – SQL, python, and tableau. I found my first dataset from uk government website. I took a look at the data using pandas and made a quick visualisation.  Google dataset search - https://datasetsearch.research.google.com/  Keywords – ‘UK housing 2023’  Used this dataset - https://www.gov.uk/government/statistical-data-sets/uk-house-price-index-data-downloads-may-2023  License - UK HPI data is published under Open Government Licence. When using or publishing data from the UK HPI reports, background tables in the statistical datatset: UK House Price Index: data downloads or search tool, you will need to add the following attribution statement:  Contains HM Land Registry data © Crown copyright and database right 2020. This data is licensed under the Open Government Licence v3.0  I believe the data to be of high quality after reading the information about where it is sourced from here - https://www.gov.uk/government/publications/about-the-uk-house-price-index/about-the-uk-house-price-index#the-difference-between-the-uk-hpi-and-others. It is updated regularly. |
| **Other thoughts or questions:** | Interesting data from Northern Ireland. Not sure what region of UK I want to focus on. Need to see what regions of other data I can locate first. |

|  |  |
| --- | --- |
| **Date:** 1/2/24 | **Course/topic: Second Dataset** |
| **Prompt:** | Median UK wages  Spreadsheets |
| **Journal Entry:** | To give my housing prices more context the first thing I want to do is compare it to the average UK salary. After searching for appropriate datasets, I found this from the Office for National Statistics. It was only available as an excel sheet. I decided to work with the data in Excel before saving it as a csv to import into my notebook. It lists the percentiles for weekly wages in the UK from 1968 to 2023. It is different to my other dataset in that they only have 1 reading per year. 1969 is missing. 1968-1983 are split by male and female. The rest are the average of both genders. See my changelog document for information on how I cleaned this data for use in my notebook.  Searched on google datasets keywords ‘UK earnings 2023’  https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/earningstimeseriesofmediangrossweeklyearningsfrom1968to2022  Again, licensed under Open Government license I have added statement to the bottom of my notebook and will do the same to my finished article. |
| **Other thoughts or questions:** | Once visualising it clearly shows a massive increase of the cost of a home compared to the annual salary. This is compelling data to argue that the difficulty of buying a home has increased.  The data is of high quality, and I checked their sourcing methods here - https://www.ons.gov.uk/methodology. It is updated regularly. |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

|  |  |
| --- | --- |
| **Date:** | **Course/topic:** |
| **Prompt:** |  |
| **Journal Entry:** |  |
| **Other thoughts or questions:** |  |

### Need another journal entry template?

If you have more journal entries to complete, please copy and paste one of the tables above and use it as a template for future entries.