**Project diary** [Copy and paste the box for each week. No more than one A4-page per week.]

|  |  |
| --- | --- |
| **Week** | 1 (Week begin 30/03/2020) |
| **Communication** | * Fiona made original contact with the group (28/03/2020) * All but two in the group responded by 01/04/2020 * Amit suggested we set a time and date to discuss and divide the work and Fiona suggested we do this by the 01/04/2020. * Fergal and Fiona discussed their ideas and progress via Microsoft teams (04/04/2020). Vishal and Amit contacted the group later, with Vishal proposing work to be done in Tableau. |
| **Any analysis or other difficulties that arose and how they were overcome** | * Finding contact details of members of the group * Fergal had issues parsing unicode into R when loading as dplyr tibble, solution was to load into R as dataframe first and then parse. * Fiona had issues reading in the data due to euro symbols. Solution was to read in without header=TRUE, add new column headings to correspond with Fergal’s and removing the now defunct row 1 (the original headers). |
| **Progress made** | * Fergal set up two github repos and a project WhatsApp group. * Fergal pushed initial .csv data to repos, added basic parsing for Price column to convert to numeric and began investigating Chloropleth Maps for representing data. * Fiona starting looking into the individual towns in more detail. She decided on 16 towns across Ireland and extracted the data for these and made a preliminary start on visualisation of the data. |
| **Member contributions to progress** | * Chloropleth Map code - Fergal * Town code - Fiona |

|  |  |
| --- | --- |
| **Week** | 2 (Week begin 06/04/2020) |
| **Communication** | * Discussed with team members using Tableau as well for visualisation. * Meeting was held on the 11th April. Fiona suggested that we should all work off the one clean dataset, and offered to make a start on cleaning the data once this was agreed by the team members present. She also pointed out that the project description asks us to look at house price change over a range of Irish towns as opposed to counties. * A deadline of the 22nd May was set at the meeting to have a draft report complete. * A second meeting was held on 12th April. |
| **Any analysis or other difficulties that arose and how they were overcome** | * Rgeo package required rebooting rstudio in order to work. * Multi-unit complexes are recorded in quite a variety of ways. Fiona contacted the group to discuss how we were to deal with this (via WhatsApp) and we decided to drop any properties over 10 million euro. Fergal came up with some code to spilt multi-property units into individual units – with an imputed mean house price. This was added to the code Fiona was working on to clean the dataset. * For the town dataset that Fiona was working on some of the house addresses are in Irish so fadas were throwing out error messages. Fiona used str\_replace\_all(hp$address,"[^[:alnum:]]", " ") to get around this issue. |
| **Progress made** | * Fergal sourced .shp file to use for chloropleth map, loaded .shp file into r and plotted basic map of Ireland (without analysis) * Fiona has looked at median house price change over time for 16 different towns around Ireland. * Fiona and Fergal worked on cleaning the dataset and uploaded it to github for the rest of the team to check on 12th April. |
| **Member contributions to progress** | * Fiona: cleaned dataset with aid of Fergal. Produced code to extract data on 16 towns around Ireland to work with. |

|  |  |
| --- | --- |
| **Week** | 3 (Week begin 13/04/2020) |
| **Communication** | * A large number of visualisations from members of the team were presented to the rest of the group via email. * Meeting held on 15th April. * We agreed that each member would have half a page to present their visualisations along with supporting test. The final half page would be used for our conclusions. |
| **Any analysis or other difficulties that arose and how they were overcome** | * Duplicate properties within the dataset. Some of these are due to the same property being sold multiple times throughout the timeframe, some are an artefact of splitting the multi-unit properties when cleaning the dataset and some are how they were entered onto the PRR to being with. For the last category, we have no way of knowing which are errors and which are multi-unit properties all sold under the one address. We therefore left them in. |
| **Progress made** | * All team members are firming up on their visualisations |
| **Member contributions to progress** | * John: heat-map displaying house price change for Dublin commuter belt * Fergal: choropleth maps * Ramesh: new house and second-hand house variation across Dublin and other towns * Fiona: visualisation on house price change over time for nine different regions in Ireland, based on four towns per region |

|  |  |
| --- | --- |
| **Week** | 4 (Week begin 20/04/2020) |
| **Communication** | * Item 1 * Item 2 * Etc |
| **Any analysis or other difficulties that arose and how they were overcome** | * Item 1 * Item 2 * Etc |
| **Progress made** | * Item 1 * Item 2 * Etc |
| **Member contributions to progress** | * Item 1 * Item 2 * Etc |

|  |  |
| --- | --- |
| **Week** | 5 (Week begin 27/04/2020) |
| **Communication** | * Item 1 * Item 2 * Etc |
| **Any analysis or other difficulties that arose and how they were overcome** | * Item 1 * Item 2 * Etc |
| **Progress made** | * Item 1 * Item 2 * Etc |
| **Member contributions to progress** | * Item 1 * Item 2 * Etc |