COMP20008
Final Report - Group 13

INTRODUCTION

Climate change and the environmental ramifications of non-renewable energy are some of the great challenges of our time, as is growing inequity within societies. Governments have a key role to play in addressing such issues, and the Australian experience with rooftop solar makes for an interesting case study combining these topics.

There are both federal and state-level incentives available to encourage solar panel adoption amongst the general populace. The combined impact of these schemes in Victoria, along with associated costs currently standing at less than half of those a decade ago[1], entails a payback period of between 3 to 6 years[2]. Victorian solar rebates have paid over $237 million towards panel installations, with the average Victorian household saving approximately $890 per year on energy bills by installing solar, for total annual savings of over $149 million[3]. Such savings would have an especially large impact on the household income of lower income groups, but are these groups seizing the opportunity?

By analysing which income groups have most been utilising solar rebates over almost two decades, this research project aims to determine the accessibility of the initiative across varying socioeconomic classes, using public data. This addresses the liveability, inclusiveness and sustainability of communities in Victoria.

DATA WRANGLING

In our analysis we elected to use Victorian data only, as including other states would necessitate trying to split out the impact of the various types of state-level incentives available.

There were four datasets used to investigate the research aims:

Description	Source	Format	Size (MB)
Monthly solar installation rates, by postcode	• http://www.cleanenergyregulator.gov.au/RET/Forms-and-resources/Postcode-data-for-small-scale-installations/historical-postcode-data-for-small-scale-installations • http://www.cleanenergyregulator.gov.au/RET/Forms-and-resources/Postcode-data-for-small-scale-installations#Postcode-data-files	CSV	11
Free-standing and semi-detached dwellings census data, by postcode	https://auth.censusdata.abs.gov.au/webapi/jsf/login.xhtml	CSV	0.1
Household income, by postcode	https://data.gov.au/data/dataset/taxation-statistics-postcode-data	CSV	60
Spatial Geometry for each postcode	https://spatialvision.com.au/blog-open-source-spatial-geopandas-part-1/	SHP	71

[1] https://www.solarchoice.net.au/blog/solar-power-system-prices/
[2] https://www.energycouncil.com.au/media/jv4blk2l/final-pdf-australian-energy-council-solar-report_-jan-2021.pdf
[3] https://www.solar.vic.gov.au/solar-homes-program-reporting

Below are the attributes we extracted for analysis, each of which was compiled per postcode, per year:

Solar installation data is published annually:

1) Total number of solar installations

The census dwelling data was extracted using the Australian Bureau of Statistics table builder app:

2) Total dwellings
3) Viable dwellings (number of dwellings where it would be feasible for a single household to install solar, i.e. free-standing and semi-detached houses)

The census data for dwelling types was available for 2006, 2011 and 2016. Other years were extrapolated / interpolated according to simple linear building growth rates between these years.

The Australian Taxation Office household income data was delivered in a highly variable format, so some wrangling was required to pull back consistent data:

4) Total household income

Postcode and year were incorporated together as a composite primary key for joining the datasets. We subsequently compiled two key metrics that were utilised throughout the analysis:

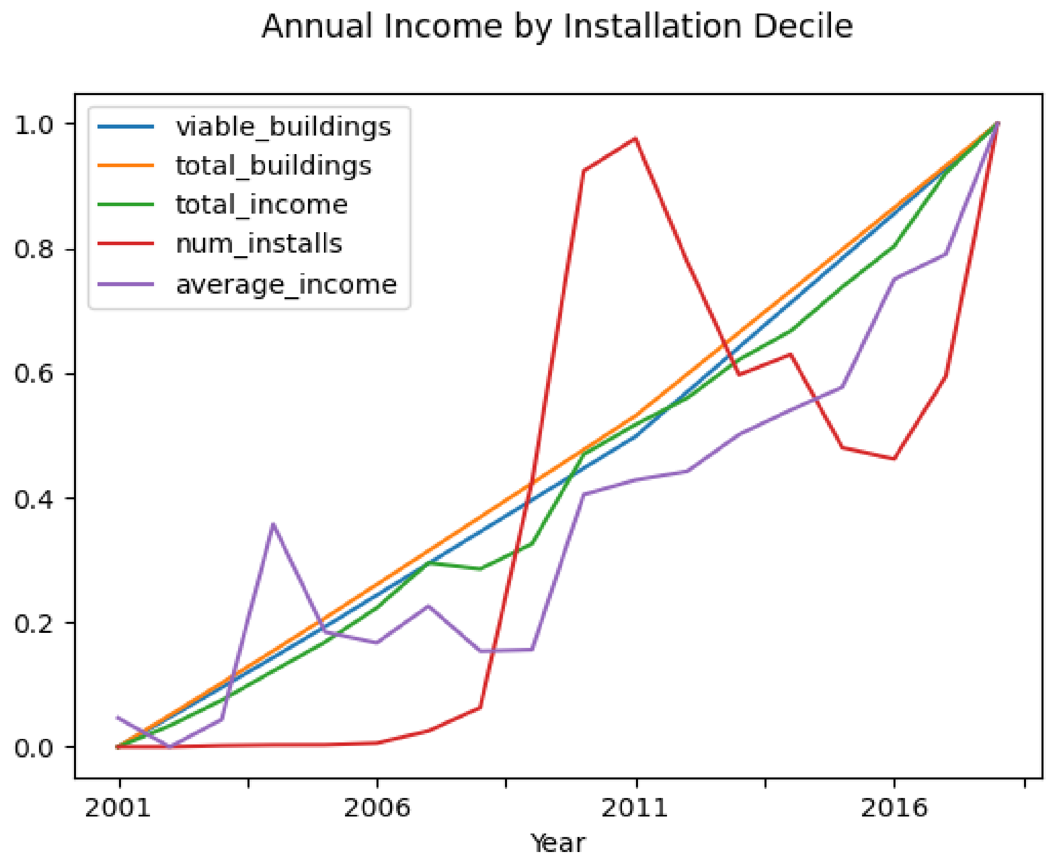
5) Average household income = total income per postcode / total dwellings
6) Solar penetration = total number of solar installations per postcode / total number of viable buildings

Ideally we would have data showing the exact income of every household that installed solar in a particular year, which could readily be used to demonstrate whether solar installation was equitably distributed across different income groups. This data is not available, however, and thus we instead took a postcode-level view of household income. We assume, for instance, that if a lower income postcode exhibited high installation rates then, on average, it was lower income households installing solar.

ANALYSIS AND RESULTS

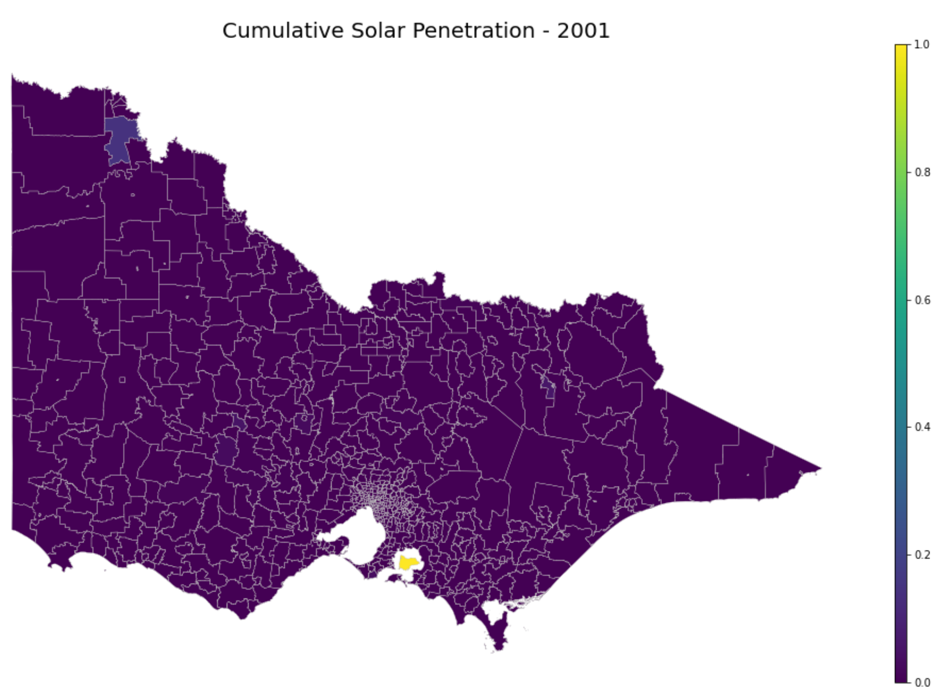
Overview

Before further analysis we examined the Min-Max normalised growth rate of the key attributes below. It can be seen that dwelling growth was consistent, income growth showed some variation, whilst solar installation rates were highly variable, with very few installations in the first 6-7 years.

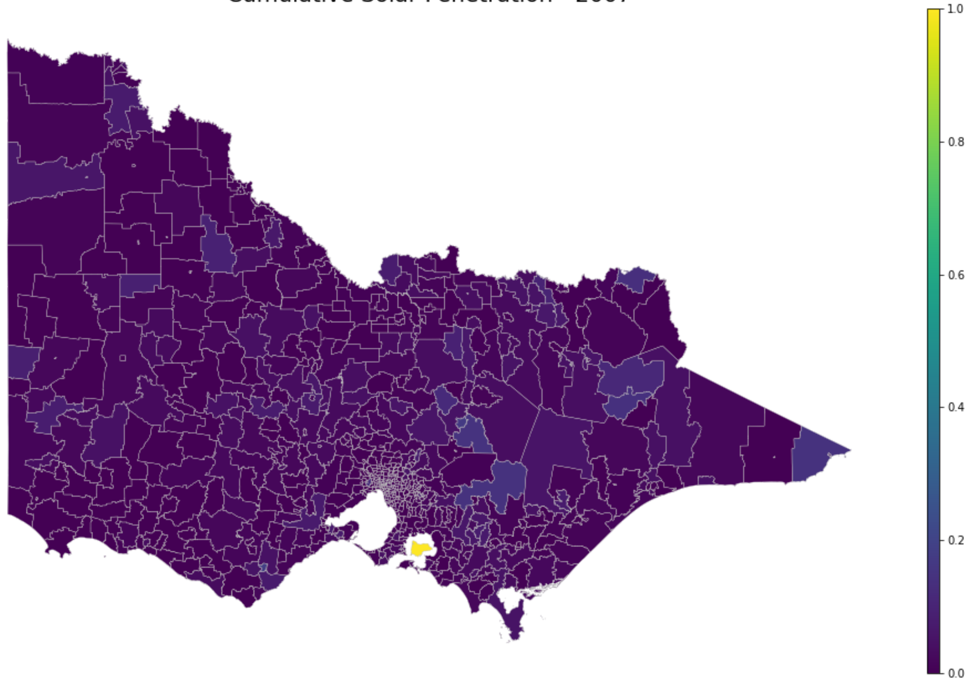


Geographic Heatmaps

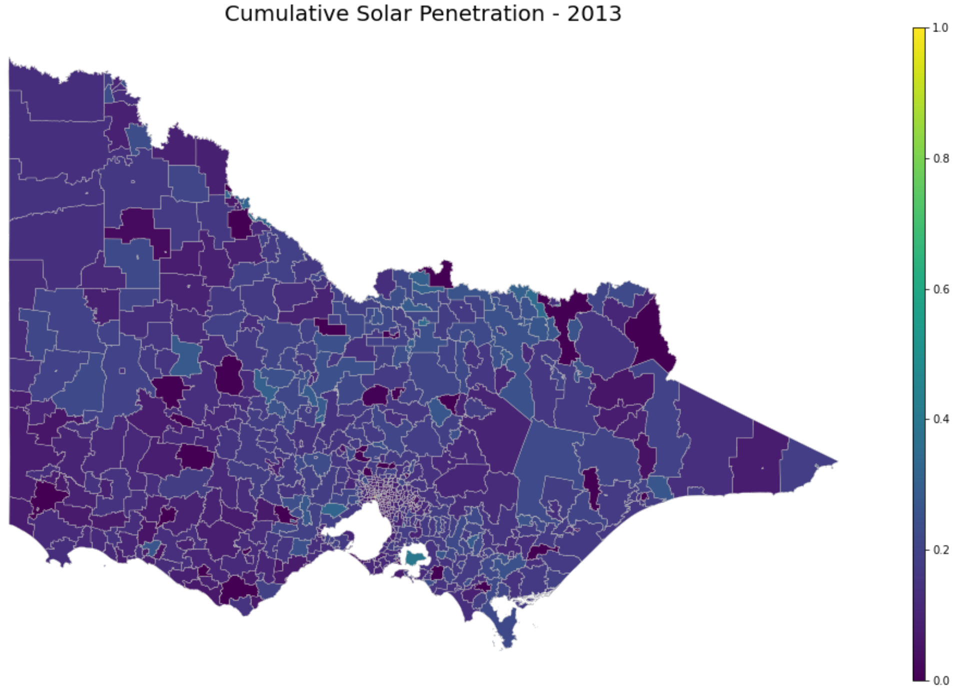
Geographic heatmaps were created to visualise the growth and spread of solar penetration across different Victorian postcodes, as well as the evolution in average incomes. A base table containing the Victoria postcodes mapped with its respective spatial geometry was used to implement this. Some wrangling was required to manage outliers, as all the postcodes required values to be graphically presented.



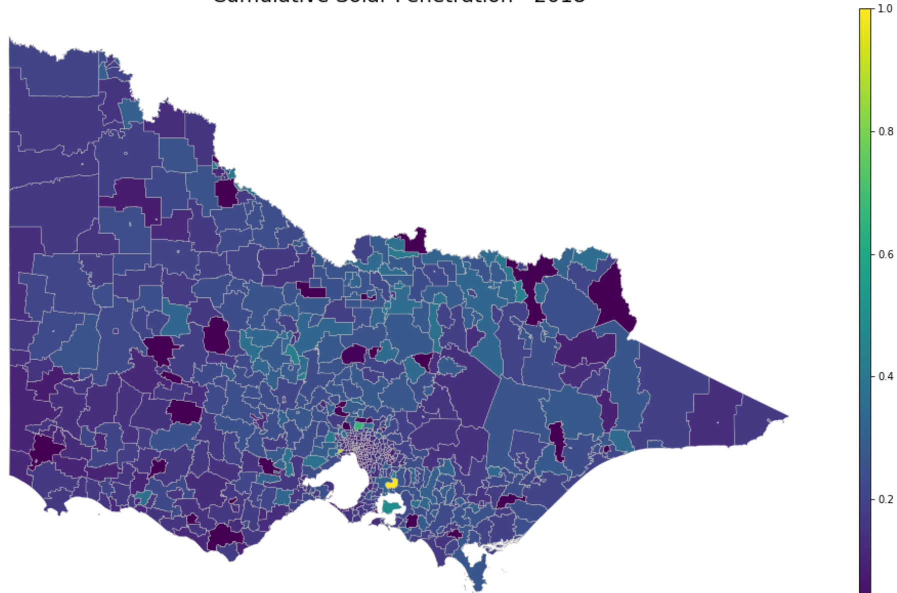
Cumulative Solar Penetration - 2007



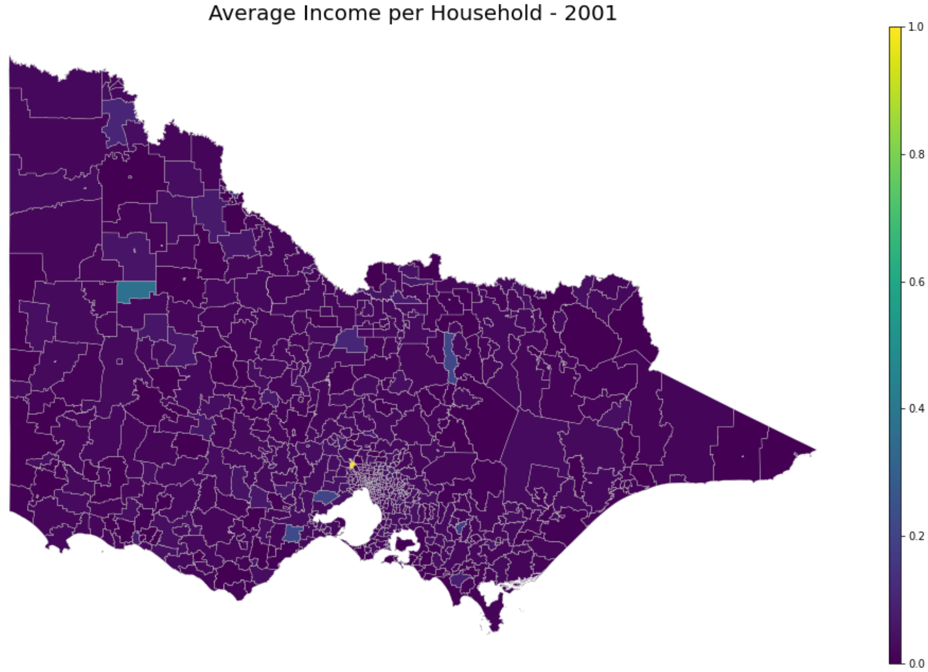
Cumulative Solar Penetration - 2013



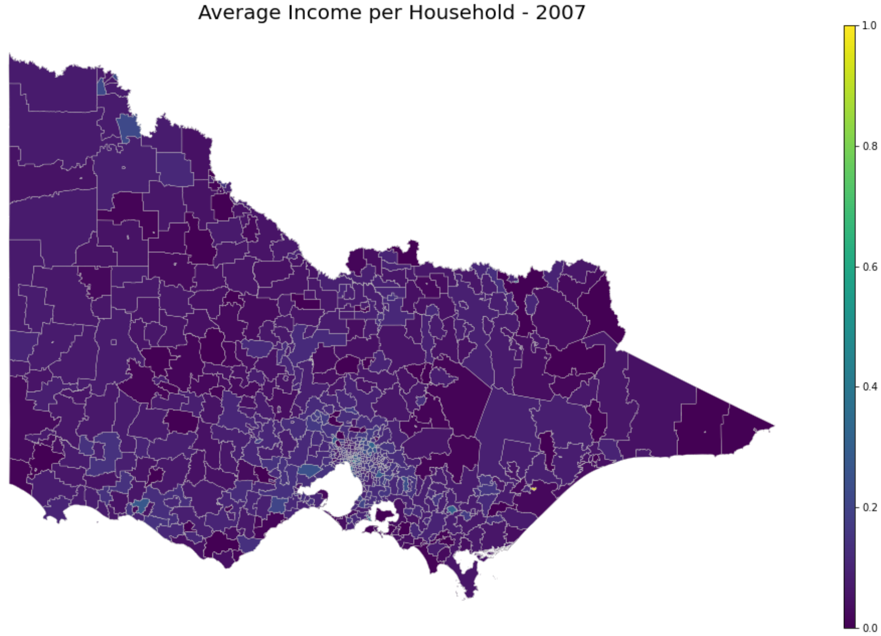
Cumulative Solar Penetration - 2018



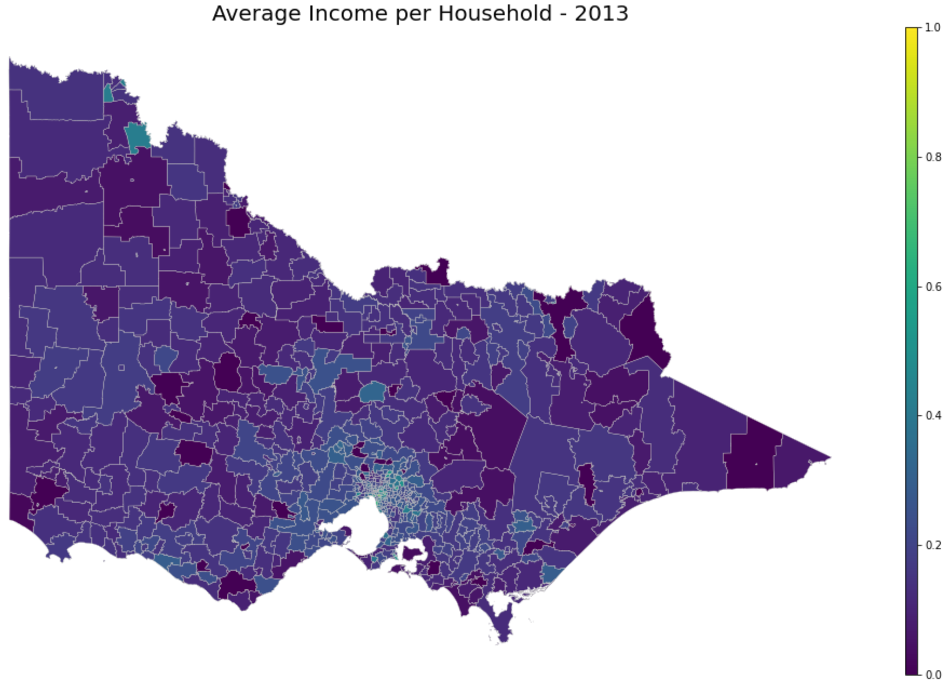
Average Income per Household - 2001

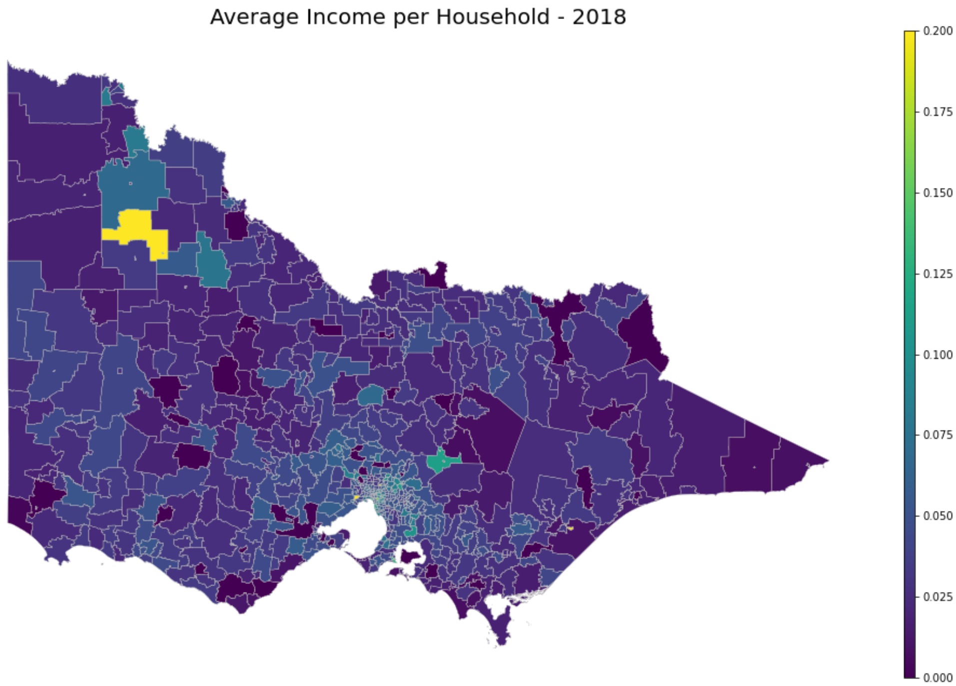


Average Income per Household - 2007



Average Income per Household - 2013





Note that there may be some cases of missing data where we observed missing installation rates and average income. In earlier years a large disparity can be seen in installation rates across postcodes, reducing over time, with the central region of Victoria seeming to have grown the most. A similar trend is also visible for the average income per household, with income distribution becoming more evenly spread across the community. The following is a sample of a few postcodes that demonstrate the typical trajectory seen for annual solar penetration:

Postcode - 3978	2001	2007	2013	2018
Installations	0	0	66	1043
Viable buildings	415	438	1802	5174
Install Percent	0.00%	0.00%	3.66%	20.16%

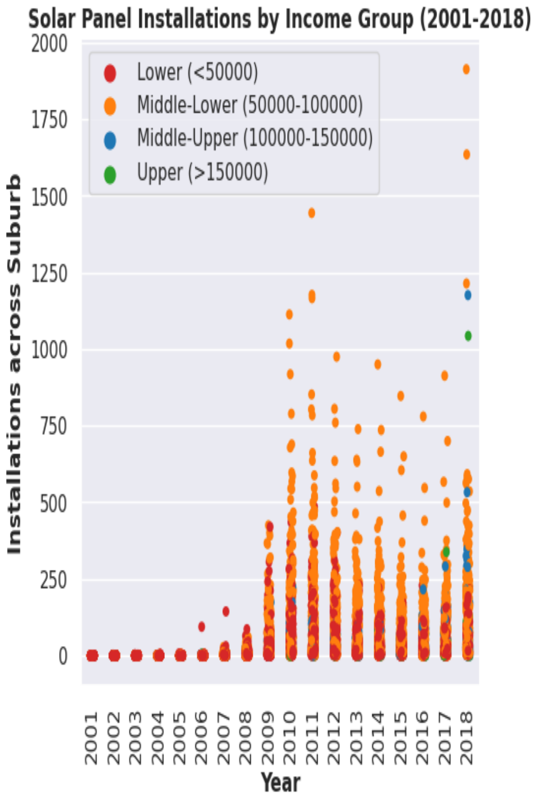
Postcode - 3746	2001	2007	2013	2018
Installations	0	0	8	14
Viable buildings	103	116	144	192
Install Percent	0.00%	0.00%	5.56%	7.29%

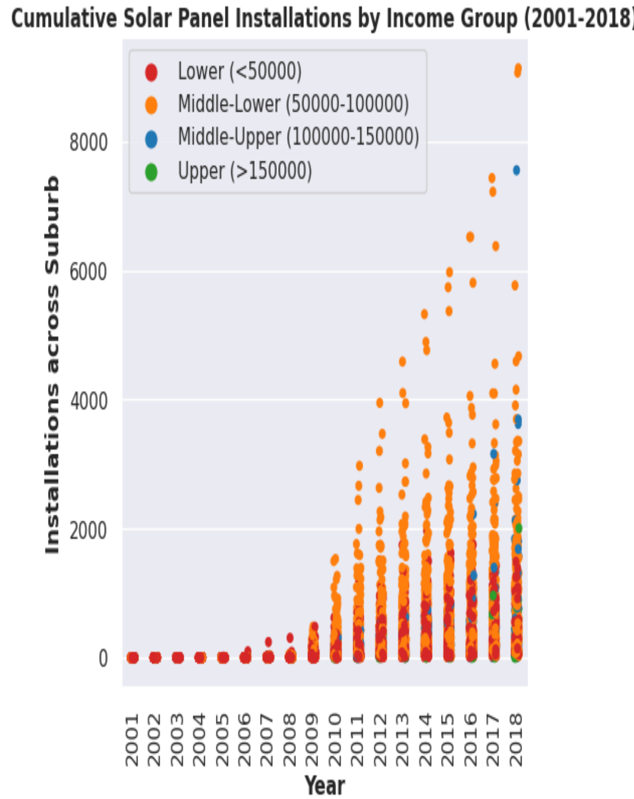
Postcode - 3750	2001	2007	2013	2018
Installations	0	0	56	313
Viable buildings	77	155	1306	4058
Install Percent	0.00%	0.00%	4.29%	7.71%

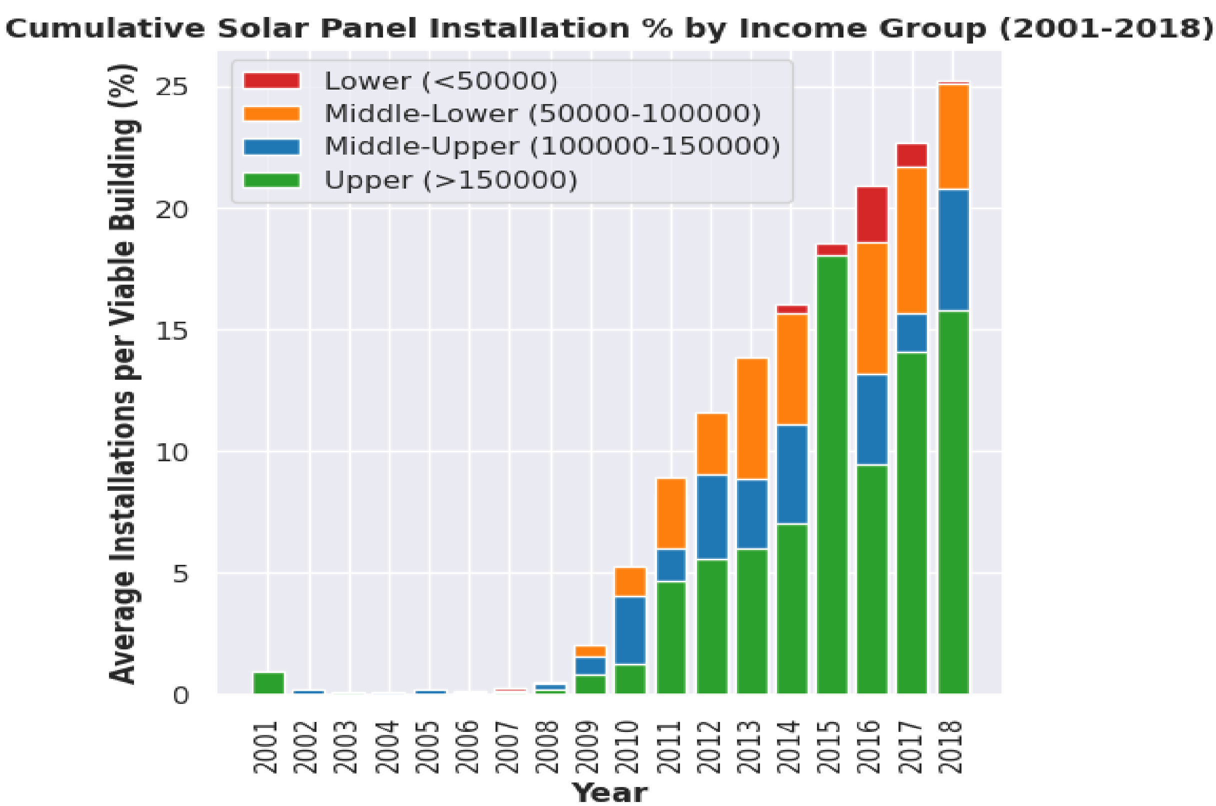
Postcode - 3975	2001	2007	2013	2018
Installations	0	1	165	289
Viable buildings	103	2123	4003	4662
Install Percent	0.00%	0.05%	4.12%	6.20%

Income Bands

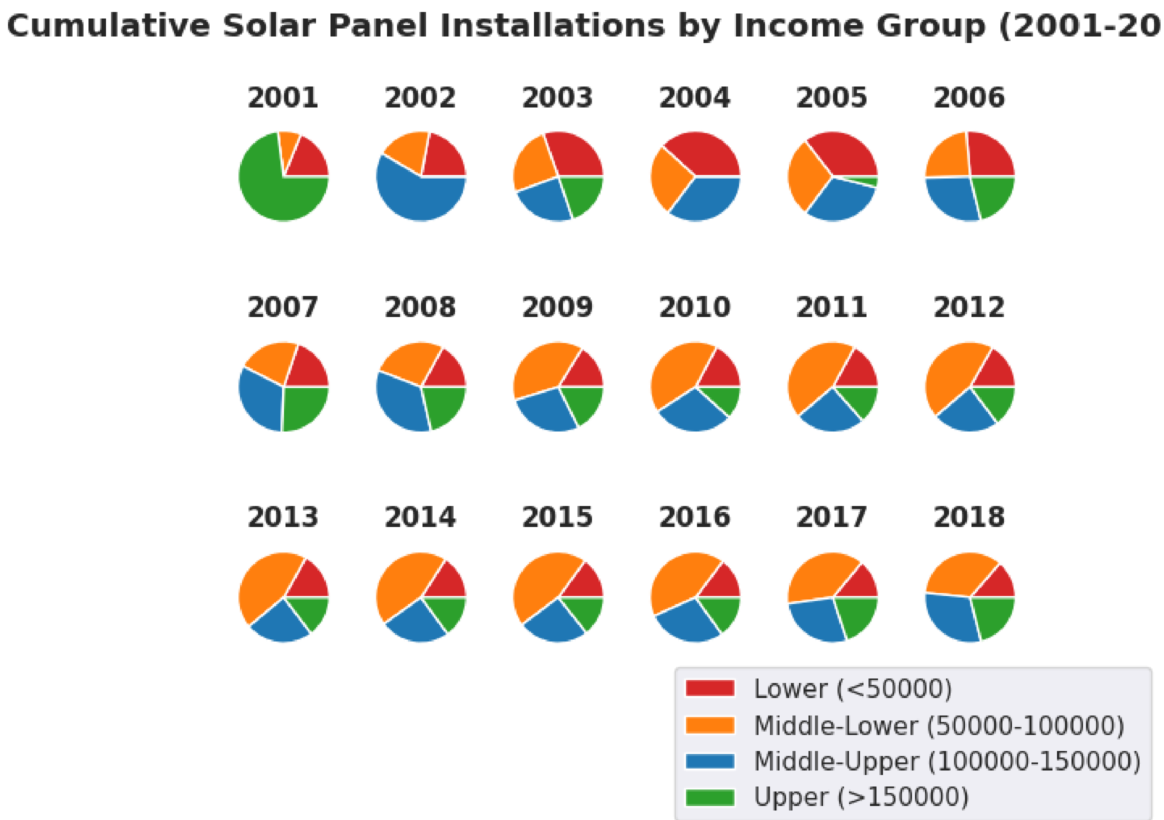
Our next step was to partition postcodes into four static income categories: “Lower” (less than 50,000), “Middle-Lower” (between 50,000 and 100,000), “Middle-Upper” (between 100,000 and 150,000) and “Upper” (greater than 150,000).


Solar Panel Installations by Income Group (2001-2018)
Lower (<50000)
Middle-Lower (50000-100000)
Middle-Upper (100000-150000)
Upper (>150000)
Installations across Suburb
Year


Cumulative Solar Panel Installations by Income Group (2001-2018)
Lower (<50000)
Middle-Lower (50000-100000)
Middle-Upper (100000-150000)
Upper (>150000)
Installations across Suburb
Year


Cumulative Solar Panel Installation % by Income Group (2001-2018)
Lower (<50000)
Middle-Lower (50000-100000)
Middle-Upper (100000-150000)
Upper (>150000)
Average Installations per Viable Building (%)
Year

The lowest income postcodes executed the least number of panel installations, with none having annually installed more than 500 and the overwhelming majority totalling approximately 200 during any given year. Postcodes corresponding to the second-lowest income group comprise the majority of variance, with a number of points located away from the main range of 0-500. The two highest income groups, furthermore, were heavily scattered across years without producing discernible clusters. (Note that the lowest income group is underrepresented on the percentage histogram due to a high number of lower income postcodes with little to no installations resulting in lower percentages).

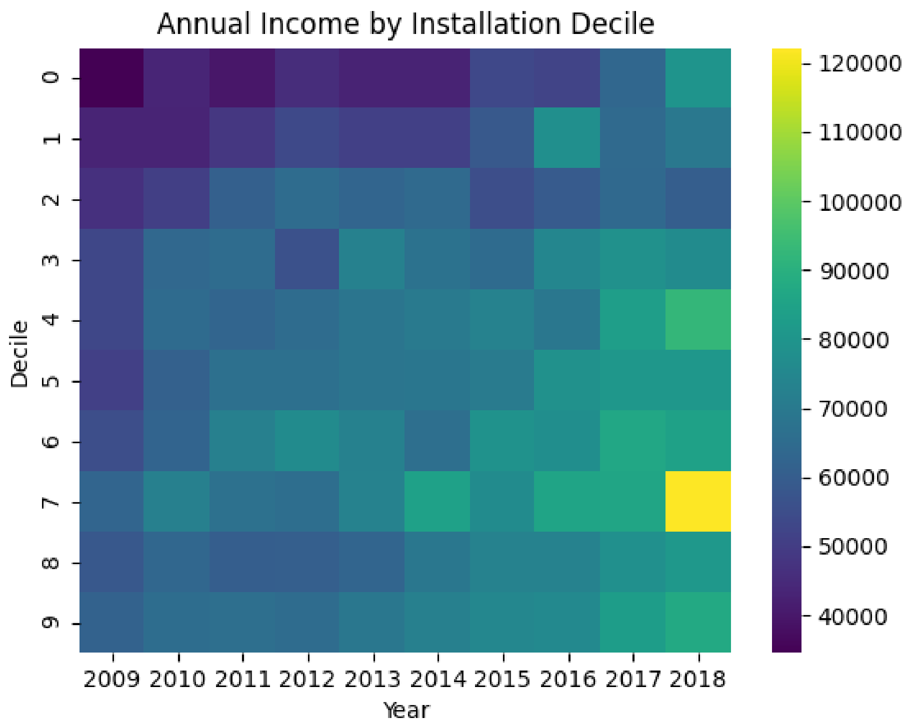


Whilst during the first year the upper income group was responsible for the majority of solar installations, in the following years the proportion attributed to this group fluctuated meaningfully. The lowest income group was also dominant in the first six years, where it represented approximately a quarter to a third of the corresponding chart, before diminishing in weight. Meanwhile, the two innermost income groups together constituted approximately two thirds of distribution for the years until 2008 (excluding the first), after which the proportion associated with Middle-Lower gradually increased until it comprised almost half of the graph until 2016, before both innermost income groups remained of equal size for the final years.

Installation Bands

In this section we reversed the analysis, and broke up the annual data into solar installation deciles, analysing the average income within these deciles. The procedure was as follows:

1) Postcodes were ranked for each year by the number of installations.
2) Ranked postcodes were divided into equal deciles, with postcodes recording no installations being excluded so as to ensure as even a distribution as possible.
3) Average income composition was analysed over time across each installation decile, and visualised as a heatmap. Data from 2001 to 2008 was excluded as there was not enough installation number variance to properly divide the postcodes into deciles.



The following conclusions were drawn:

- Overall average income increased over time across all installation deciles due to wage growth.
- Average income across installation deciles became more evenly distributed as time increased, with lower income groups in earlier years having a tendency to record lower installation rates.

Income and Building Quantiles

In this section we attempt to deal with the growth over time in average income, viable buildings, and number of installations, via the following procedure:

1) Each year rank the postcodes from lowest to highest average income.
2) Divide the ranked postcodes into quantiles such that there is an approximately equal stock of viable buildings in each. Effectively we broke the total stock of viable buildings into equal-sized quantiles, with each quantile representing a different average income group.
3) Adjust installation numbers per quantile to account for the fact that the deciles aren't exactly equal in terms of number of viable buildings. This is done because one metric we looked at was the share of total installation going to each quantile.

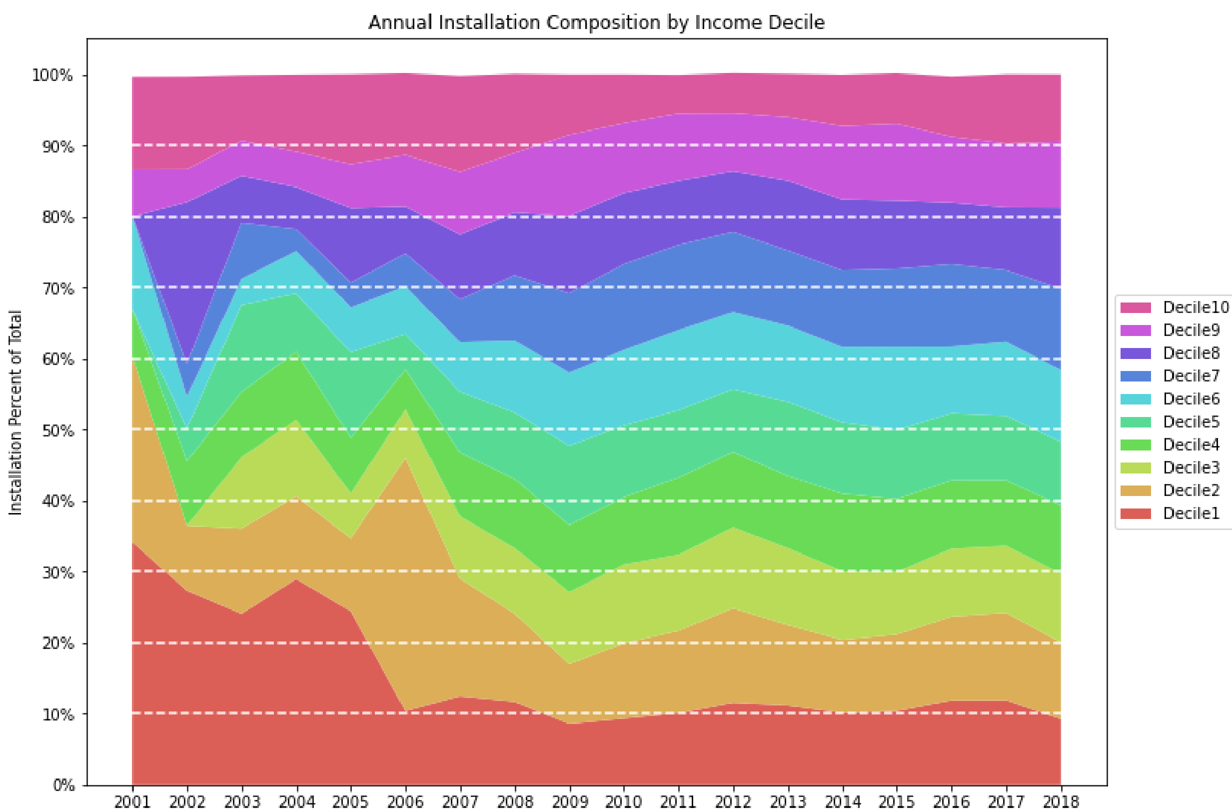
This process is illustrated for 2018 below:

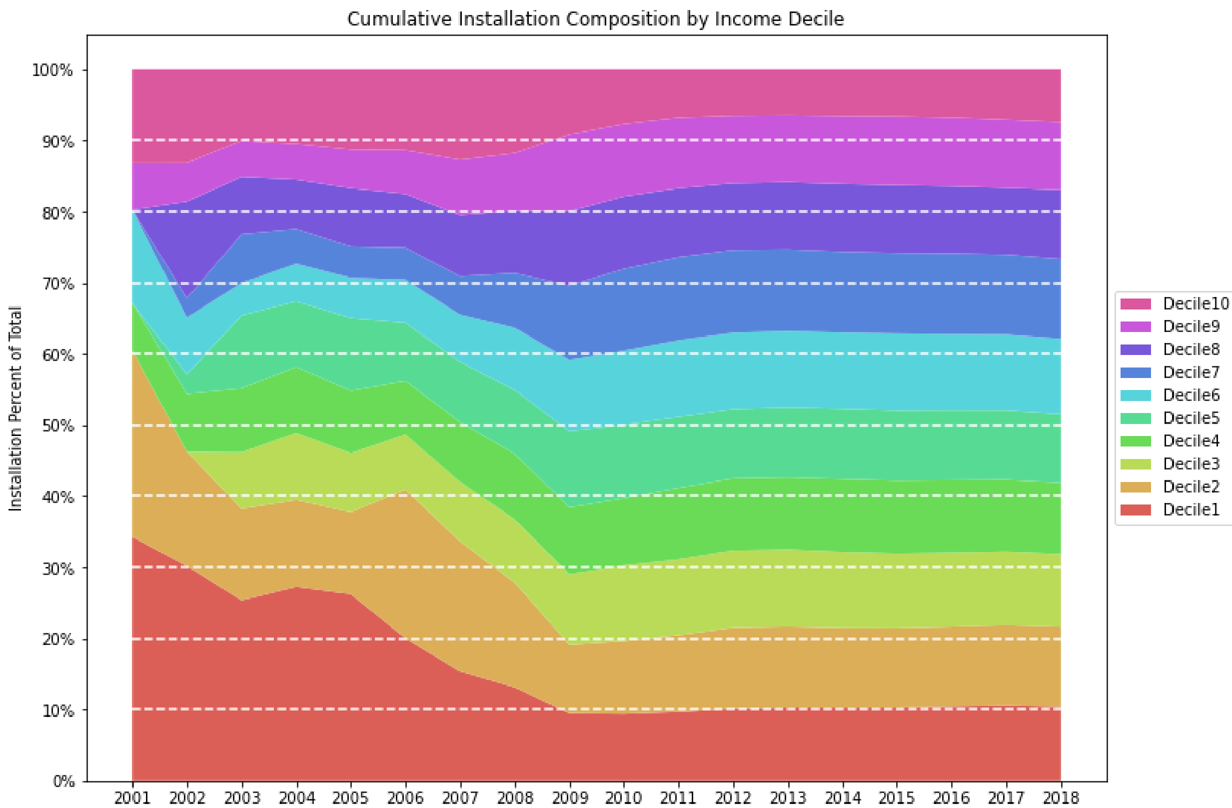
income_decile	num_postcodes	average_income	install_num	viable_buildings
1	140	41245	5402	224360
2	94	57463	6211	221204
3	68	66720	5561	220298
4	42	75359	5790	232175
5	49	82765	5090	218202
6	43	90081	6106	231593
7	29	95130	6639	222501
8	36	99194	6657	224558
9	59	110672	5389	222699
10	67	207338	5656	228822

It can be observed that there is a roughly equal share of viable buildings in each income decile. It is also interesting to note that the lowest and highest income deciles have the highest number of constituent postcodes, likely because both the lowest and highest income suburbs will both tend to have lower density than middle class suburban postcodes.

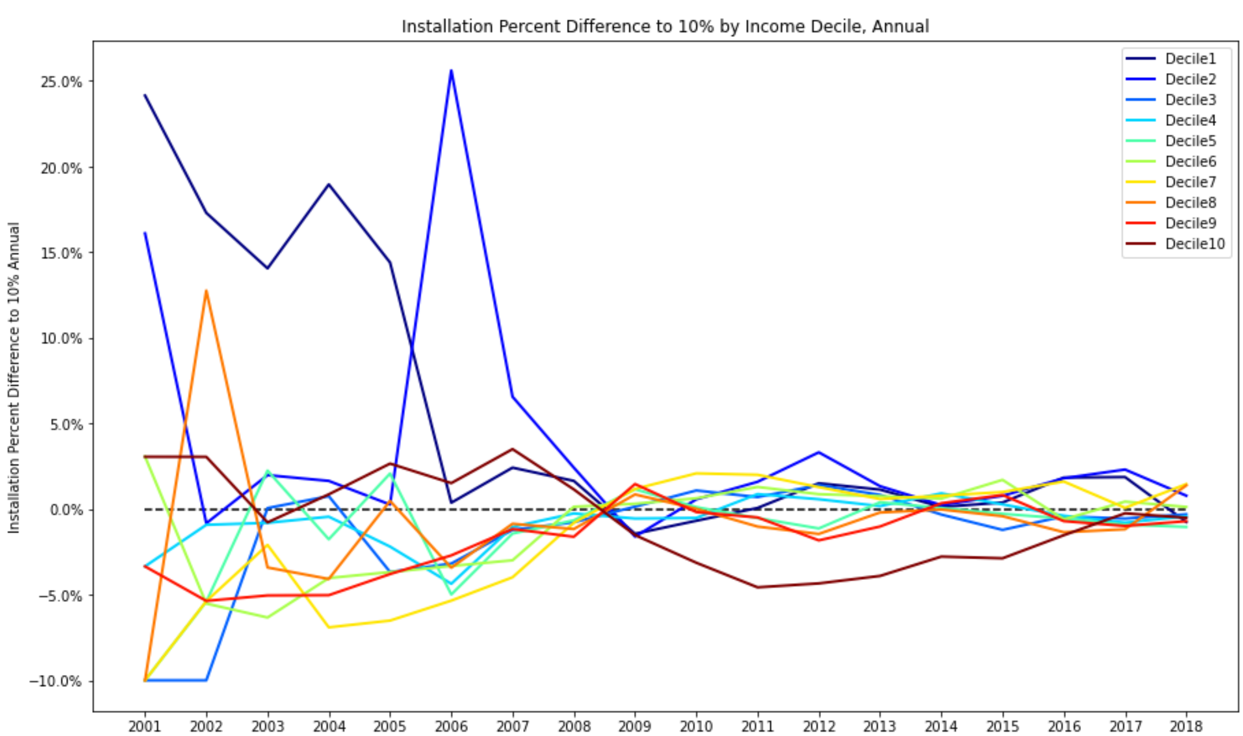
We then looked at two measures:

1) The installation share by year, both absolute and cumulative, going to each quantile (decile or quintile)
2) The difference to a "fair share" (10% for deciles, 20% for quintiles) of installations going to each quantile, both by year and cumulative
3) We show deciles for the granularity this affords, and quintiles to aid visual interpretation

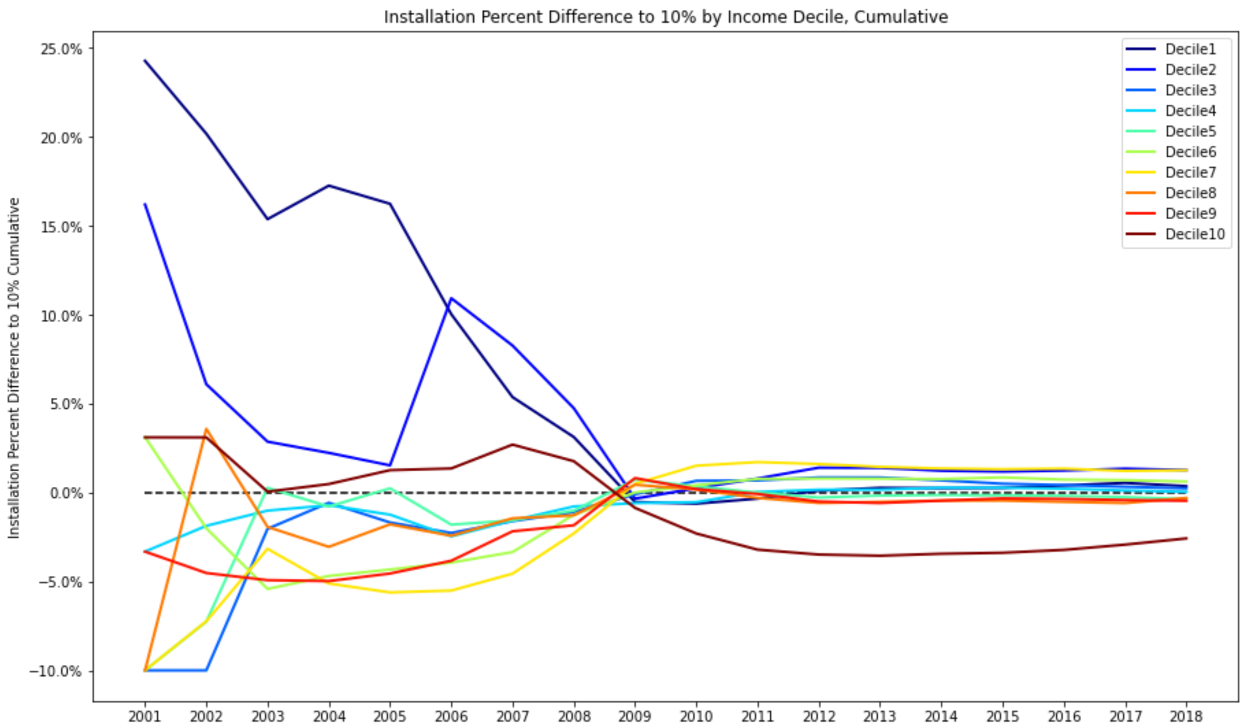


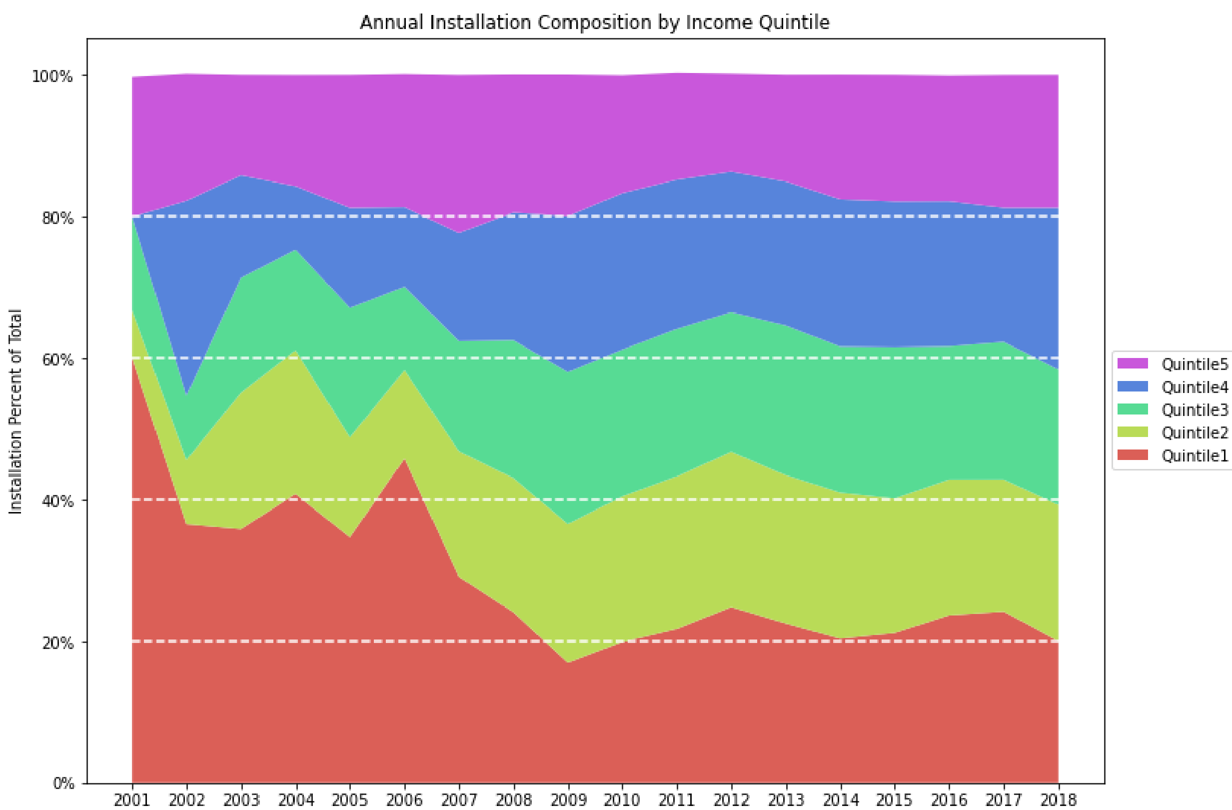


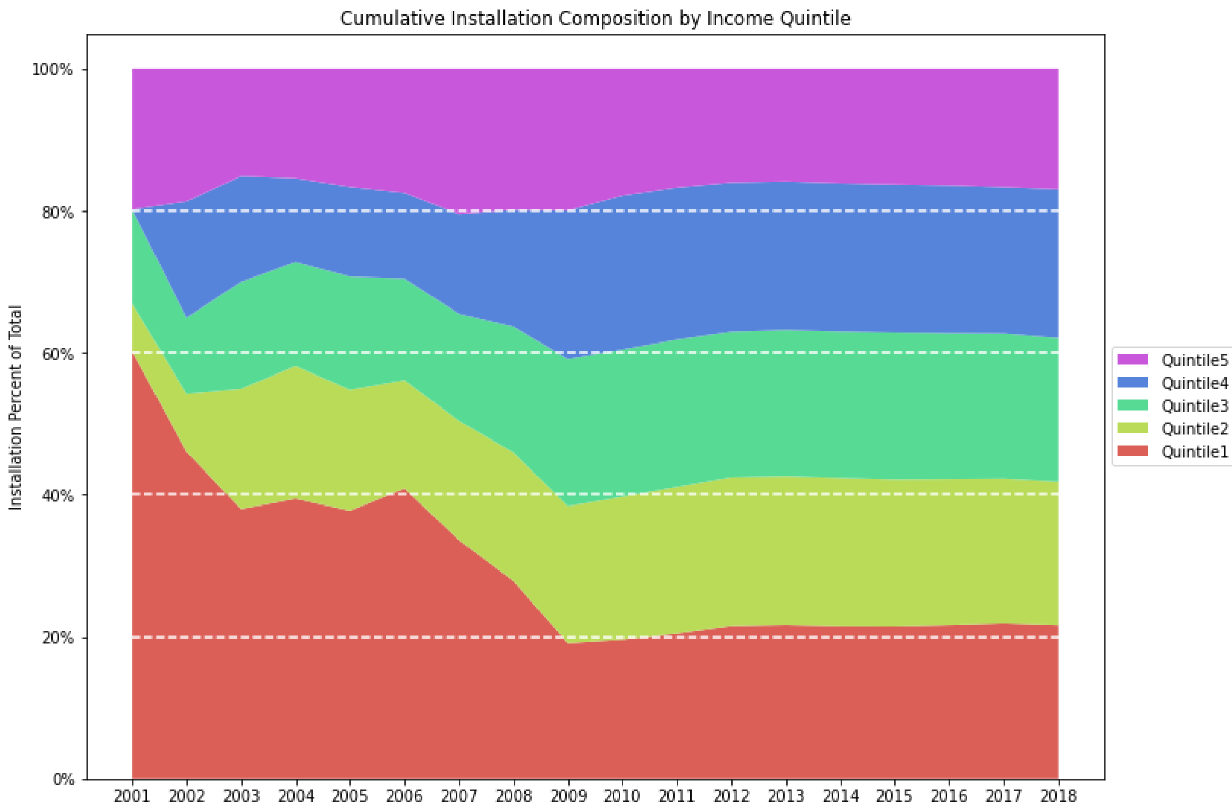
Cumulative Installation Composition by Income Decile
Installation Percent of Total
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Decile10
Decile9
Decile8
Decile7
Decile6
Decile5
Decile4
Decile3
Decile2
Decile1

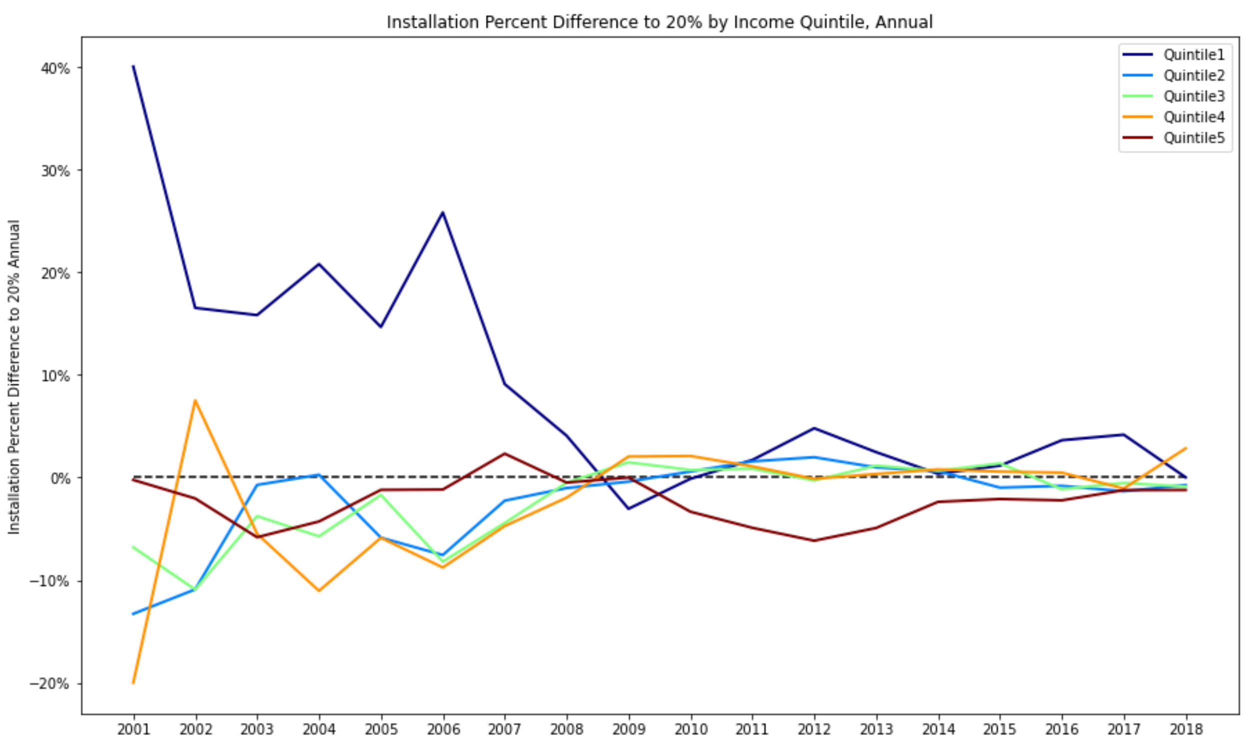


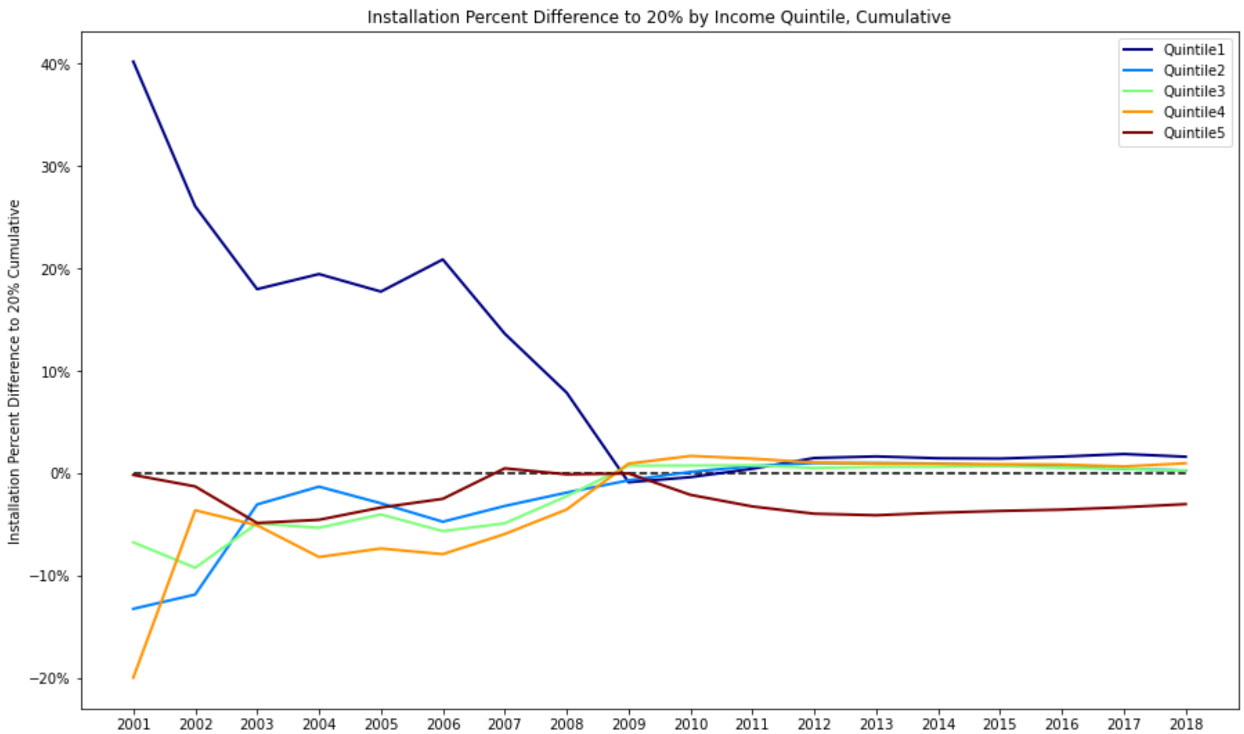
Installation Percent Difference to 10% by Income Decile, Annual
Installation Percent Difference to 10% Annual
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%
−5.0%
−10.0%
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10


Installation Percent Difference to 10% by Income Decile, Cumulative
Installation Percent Difference to 10% Cumulative
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%
−5.0%
−10.0%
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Decile1
Decile2
Decile3
Decile4
Decile5
Decile6
Decile7
Decile8
Decile9
Decile10


Annual Installation Composition by Income Quintile
Installation Percent of Total
100%
80%
60%
40%
20%
0%
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Quintile5
Quintile4
Quintile3
Quintile2
Quintile1


Cumulative Installation Composition by Income Quintile
Installation Percent of Total
100%
80%
60%
40%
20%
0%
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Quintile5
Quintile4
Quintile3
Quintile2
Quintile1


Installation Percent Difference to 20% by Income Quintile, Annual
Installation Percent Difference to 20% Annual
40%
30%
20%
10%
0%
−10%
−20%
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Quintile1
Quintile2
Quintile3
Quintile4
Quintile5



The above analysis provides several interesting conclusions:

- Contrary to earlier analysis, by this measure lower income households were in fact installing more than their "fair share" of solar in the earlier years. We do note that absolute installation levels were low for the first 6-7 years.
- Conversely, the highest income households did not dominate installations in the earlier years of the study, but instead maintained a reasonably stable share throughout.
- Middle income households were most strongly underrepresented earlier in the analysis, but over time achieved their "fair share" as well.
- From 2010 onwards the share of installations going to the different income deciles was fairly evenly distributed, pointing towards an equitable distribution of solar installation across all income brackets.
- On a cumulative basis the top income decile or quintile somewhat lagged behind the other quantiles by 2018. Perhaps for some in this group the economic motivation (in reference to earnings) was not as strong as it was for others.

We tested one hypothesis - that earlier high installation rates from lower income groups might have been driven by housing construction in newer postcodes, but found no evidence suggesting a correlation between building growth and installation rates.

All analysis methods utilised were chosen for their ability to present a reasonably large data set with two key attributes, average household income and solar penetration, that varied both across geographic region and time. We therefore experimented with different ways to encapsulate these dimensions in visualisations and analysis that were both substantive and intuitive, and feel that the variety of final approaches make for a richer understanding.

SIGNIFICANCE OF RESULTS

In this report we have taken a varied approach to analysing the research question: have the full spectrum of economic groups, as measured by household income, taken advantage of the investment and environmental opportunity in rooftop solar? Earlier analysis used a static ranking of income and this tended to show that higher income groups dominated at the beginning, with lower income groups catching up over time. However, in the income and building quantile approach, where we have effectively tried to break the viable housing stock up into equal quantiles ranked by income, we found that lower income groups in fact had an outsized share of installation in earlier years. The difference in results is due to how we have broken up the data as well as the small number of installations in early years; in the latter case we have broken up the viable housing stock in a consistent manner from year to year.

All forms of analysis certainly support the idea that middle income groups have increased solar penetration, and the contemporary distribution of solar is now equitable. Given the economic impact of rooftop solar to the household income of poorer groups, we believe that this outcome, and the overall growth in an effective climate change mitigation strategy, is something to be celebrated. We would suggest that further efforts could be made by the government, in terms of marketing and awareness campaigns, to encourage take-up by lower income groups and potentially nudge the rate of installation higher in these groups. We also suggest that such analysis as conducted throughout this report be continued and published, to ensure that solar takeup remains equitable in the future.

CONCLUSION

Limitations / Challenges

- The key limitation throughout the project was the core assumption that average household income by postcode was indicative of the average income of households installing solar panels, whereas in reality there may have been skewed distributions within postcodes.
- Census data for dwelling types was only available from 2006, 2011 and 2016, and as such building growth for other years was extrapolated according to simple linear building growth rates between three years.
- While data wrangling, we had to omit data rows containing null or missing values.
- While we analysed solar penetration by viable buildings per postcode, solar panel use may vary by postcode due to confounding variables such as level of solar irradiance and roof tilt.

Future Research

The ideal extension would be to obtain actual household income per solar installation data and use statistical methods to extrapolate this to the more general data set, however such data is not publicly available.

This project could also be further extended based on data from the National Electricity Market, an Australian Energy Market Operator, from which data sets containing energy price, demand history and daily change of average energy price can be extracted. Other influencing variables such as solar panel cost, available feed-in tariffs, rebate amount, and income / asset tests on rebates could also be considered, as could considering data from wider regions across Australia.