



# The Southern Initiative Income, Wealth and Wellbeing Analytics Project:

# Final Report

Prepared by Harmonic Analytics Ltd

for Auckland Council

### Disclaimer

The results in this report are not official statistics, they have been created for research purposes from the Integrated Data Infrastructure (IDI) managed by Statistics New Zealand.

The opinions, findings, recommendations and conclusions expressed in this are those of the author(s) not Statistics NZ.

Access to the anonymised data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business or organisation and the results in this [report, paper] have been confidentialised to protect these groups from identification.

Careful consideration has been given to the privacy, security and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the Privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz.

The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information may be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes.

Any person who has had access to the unit-record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to secrecy. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

# **Version Control**

Issue	Version	Who	Date	Comments
1	DRAFT 1	Lisa Chen, Harmonic Analytics	nic Analytics 12/6/2019 Initial Draft	
1	DRAFT 2	Shirley Wu, Harmonic Analytics	13/6/2019	Reviewed initial contents
1	DRAFT 3	Harmonic Analytics team	21/6/2019 Released for initial client feedback.	
1	DRAFT 4	Harmonic Analytics team	28/6/2019	Update content based on client's feedback.
1	Final	Harmonic Analytics team	16/7/2019	Update plots and figures

## **Executive Summary**

This document presents the key findings of The Southern Initiative (TSI) Income, Wealth and Wellbeing Analytics Project undertaken by Harmonic Analytics from January to July 2019. The findings of this project will contribute to our understanding of the relationships between education, vocation, income levels and wellbeing. This improved understanding will ultimately, help ensure a more equitable distribution of Auckland's prosperity for South and West Auckland.



By visualising the relationships between education level and annual income across various industries, it is obvious that some of the industries require less education can still abstain a reasonble yearly income. These initial findings motivated this study to identify which industries are more 'economical' (in terms of investment in education versus the income earned). Consequently, this provides clues and evidence to support policy making for improving the fairness across social welfare.

The basis of the analysis includes data from the StatsNZ Integrated Data Infrastructure (IDI) environment, including census 2013, IR 2013 and IR 2018 data. With this data, we explored the income and employment characteristics of South and West Aucklanders, alongside the rest of Auckland. Storylines were developed to summarise the findings and the key insights are listed as following:

- Generally, an individual's annual income increases with their education level. However, the rate of the increase for South Aucklanders is lower when compared with all Aucklanders, while for West Aucklanders, their rate of income increase is in line with the rest of Auckland.
- When compared to West Auckland and the rest of Auckland, South Auckland has a relatively low proportion of highly educated people.
- When considering income levels, we identified the top 5 roles for South and West Aucklanders based on an equal weight ranking on three factors, namely population count, level of education and annual income level:
  - 1. South Auckland (Industry + Occupation):
    - a. Professional/Scientific/Technical + Community & Personal Service
    - b. Transport/Postal/Warehouse + Community & Personal Service
    - c. Professional/Scientific/Technical + Labourer
    - d. Finance/Insurance + Labourer
    - e. Education/Training + Technician & Trades
  - 2. West Auckland (Industry + Occupation):

- a. Accommodation/Food + Clerical & Administrative
- b. Accommodation/Food + Professionals
- c. Agriculture/Forestry/Fish + Labourer
- d. Administration/Support + Professionals
- e. Health/SocialCare + Sales
- The Auckland population we studied in this project can be classified into four clusters: 1. Benefit recipients; 2. Pension recipients; 3. Wage and Salary earners, and; 4. Withholding Payment and Student Loan recipients.

Moreover, another 11 clusters are obtained by performing a second tier cluster analysis – i.e. cluster analysis within each of the 4 clusters. This analysis identified that age, ethnicity and employment status are the critical attributes for grouping the different populations. The analysis also found that the local board and region has no significant influence on grouping populations regardless if age and ethnicity are included in the cluster model or not. In other words, regional factor is not significant in differentiating the clusters.

• Instead of comparing the number of income sources directly, we used individual's income proportion and based on that, developed a clustering model to understand the common characteristics of sub-population with similar income composition. The cluster analysis shows that 86.5% of the South Auckland workers have a single income source. Within which, the most common income source is Wage and Salary (54%), and the most common 2-source income combination for South Auckland is Wage and Salary + Benefits (7%), followed by Wage and Salary + Pension (2%). West Auckland has a very similar pattern to South Auckland.

In addition, the age and ethnic characteristics of the benefit recipient's cluster were also studied:

- 1. 17% of the people in the *55-64 years old* age group in Auckland are classified into the cluster that are 'mostly benefit recipients', followed by 11.9% of *45-54 years old* and 10.5% of *35-44 years old*.
- 2. 22% of Middle East, Latin American or African (MELAA) in Auckland is classified as 'mostly benefit recipients', followed by 21% Maori and 19% of Pacific Island.

To reflect the temporal changes from 2013 and 2018, IR2018 data was included in the final stage analysis. Through this analysis, we found that South and West Aucklanders have experienced a relatively low increase in annual gross income compared to the rest of Auckland. In addition, the average increase of annual gross income for Maori and Pacific Island Aucklanders is only half that of other ethnicities.

A Shiny <sup>1</sup>dashboard has been developed for users to undertake an interactive, insightful exploration of Auckland's income and employment characteristics and perform 'deep-dive' analysis into the various cohorts and subgroups of interest.

Although we have made projections to 2018 values based on 2013 data, particularly around annual incomes. Future steps for this project include updating census the 2013 data with the actual census 2018 – when Stats NZ releases it. More in-depth analysis of specific questions can also be undertaken. Finally, the dashboard is highly extensible and inherently easy to adapt for other councils or organisations seeking to improve the wellbeing of New Zealanders.

<sup>&</sup>lt;sup>1</sup><u>https://shiny.rstudio.com/</u>

## Background

The Southern Initiative (TSI), within Auckland Council, has recently launched Shared Prosperity. This is a programme of work which aims to more equitably share Auckland's prosperity with South and West Aucklanders.

A key pillar of this work is "Inclusive Livelihoods". Inclusive Livelihoods seeks to tackle the inequalities in income and opportunities for South and West Aucklanders. This is about getting South and West Aucklanders set up with good quality, and future proofed jobs that provide a clear trajectory, economic security and upward mobility. Supporting people into meaningful employment and supporting the person to build from their strengths changes the way they see themselves. It unlocks assets and resources that can be mobilised to improve their lives and the lives of their Whānau.

To support this work, The Southern Initiative has commissioned research with Harmonic Analytics Limited ('Harmonic'). This research utilises the Integrated Data Infrastructure (IDI) to explore income and employment for South and West Aucklanders.

# **Analysis Objectives**

TSI Subject Matter Experts (SMEs) suggested the following two areas of focus for the analysis:

- 1. Earnings vs Education. Questions of interest could include:
  - a. Are there any (significant) differences between people's average earnings in South Auckland and other regions?
  - b. Are there any occupation/industry combinations that earn more or less than others?
- 2. Income sources. Questions of interest could include:
  - a. What is the proportion of South Aucklanders taking benefits, or having more than one income source?
  - b. What are the 'common' demographic factors amongst benefit takers in South Auckland, are these significantly different from those to the rest of Auckland?

Harmonic approached these by developing various 'storylines' to help address some of the questions raised by TSI. The rest of the section details the data used, the analysis and the insights of these storylines.

### Data

Harmonic has successfully applied for access to Statistics New Zealand's Integrated Data Infrastructure<sup>2</sup> (IDI) for this analysis. IDI holds microdata about people and households.

### **Data Sources**

Two IDI tables have been used as the primary source of information for the analysis.

Table 1: IDI tables used in the analysis

IDI Table	Description

<sup>2</sup> <u>https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/</u>

Census2013	Provides a snapshot of every person's demographic, income, education and working information such as occupation and industry
IR	Provides time-series based information regarding individual monthly earnings and taxable amount by various income types such as 'W&S' (wage & salary), 'BEN' (benefit), 'PEN' (pension), 'STU' (student- allowance), 'PPL' (parental leave), 'CLM' (ACC claims) and 'WHP' (withholding payment). We used 2013 values when consolidating with the Census 2013 demongraphics attributes. However, we used 2018 values to quantify and project the increase in income levels in the absence of Census 2018.

The pre-processing of the data consists of the following steps:

- 1. Clean the data sets. For example, apply treatments for the missing data values.
- 2. Consolidate and join tables. Mostly join Census and IR data to provide a complete view of income.
- 3. Transform and apply to geocode. This enables us to present the data in a map for more intuitive comparison and interpretation.

### **Data Quality Issues**

Although Harmonic has experience working with IDI data through previous projects, the updated data structure within IDI, along with new/updated data sources have presented new challenges here.

While IDI microdata provides a rich source of information, we have encountered a few data quality issues in this project:

Data Quality Issue	Examples	Assumptions/Decisions made to support analysis
<ol> <li>Unmatchable primary keys when joining two data sets within IDI</li> </ol>	For instance, the IR data has 32.2% of match-able individuals to the census2013 data amongst the Auckland working population.	We used individuals from the IR data set as the basis for any analysis related to income. As a result, the total working population with an IR record (in 2013) in Auckland is ~800K.
2. Ambiguous attribute definition	For instance, the 'income_source' attribution from the IR data has a value called 'BEN', i.e. benefits. However, it is not clear what benefit types this value includes; e.g. does it include the entrepreneur/start-up benefit?	We relied on the 'income_source' attribute from the IR data for ongoing analysis as IR values are systematically recorded with better completeness; whereas census data have more subjective biases (as people filled in their census).
3. Conflicting input	For example, for the same individual	We relied on the

values	(identified as having the same 'UID'), the 'source_of_income' from the Census2013 data has conflicting values with what is recorded in the IR data under the attribute 'income_source'.	'income_source' attribute from the IR data for ongoing analysis as IR values are systematically recorded with better completeness; whereas census data have more subjective biases (as people filled in their census).
4. Out-dated datasets	The census data is from 2013. The 2018 census is still being processed by Stats NZ and was not ready at project commencement.	Until the 2018 census data becomes available, using 2013 data is sufficient for this analysis. The 2013 statistics are reflected in the analysis findings. As a result, e.g. income analysis is based on the 2013 earnings.

### Limitations

Due to the confidentiality restrictions of IDI, the data released by IDI will lose a certain level of granularity. For example, the statistics can only be calculated, aggregated and presented up to the Local Board level.

Consequently, the conclusions from the analysis results would be affected. For example, if we are analysing to identify the (significant) difference in mean earnings between two industries, the analysis based on the individual microdata (i.e. UID) level could be vastly different from that of the aggregated meshblock or Local Board level. One may conclude that the difference is 'significant' based on the UID results, but cannot say the same with the aggregated outputs.

# Analysis and Insights

This section details the analysis and insights of the 'storylines' developed by Harmonic to answer TSI's questions of interest.

### Earnings (Based on Wage & Salary) vs Education

Before conducting the analysis, we first restricted our data set to the full-time working population in Auckland, i.e. the (sub) population who are usual residents aged 15+ and works 35+ hours per week. Also, we classified earnings as earnings based on wage and salary only. This was to eliminate the bias from the other types such as benefit and pension.

Apart from demographic variables such as age group and ethnicity, the key attributes we considered were:

- 1. Total gross wage & salary income p.a.
- 2. Level of education. E.g. level 1 is year 11, level 3 is year 13 (equivalent to high-school qualification), level 8 is Honours/Postgraduate Diploma, and level 9 is Master. Note that we used this as a proxy for the 'years of education' in this analysis.
- 3. Role description, which consists of two pieces of information:

- a. Occupation. E.g. 'manager' or 'labour'.
- b. Industry. E.g. 'Mining', 'Transportation' or 'Manufacturing'.

A role of an individual is described by a combination of occupation and industry, e.g. 'labour' in the 'mining' industry.

**Storyline 1:** Are there any significant differences in the earnings vs education relationship between different roles, or regions (e.g. South Auckland and West Auckland)?

### Plot A:



Figure 1 Distribution of annual income vs years of education in Auckland. This figure explains the relationship of education length and annual income of full-time wage & salary workers in Auckland. The line in green colour represents West Auckland, the red line represents South Aucklandand the blue line represents the rest of Auckland.

### Plot B:



Figure 2 Distribution of annual income vs years of education in Auckland for the top 5 most populated industries/occupations

### Key insights:

- 1. Plot A above shows that the relationship between annual income and the level of education is similar across South Auckland, West Auckland and the rest of Auckland. In general, the average income of individuals increases with their education level.
- 2. In general, the rate at which average annual income increases with education level is lower in South Auckland when compared with other regions.
- 3. Plot B shows that in South Auckland, the proportion of the highly-educated population is lower than that of West Auckland or the rest of Auckland.
- 4. The South Aucklanders who work in the Manufacturing sector have a higher average annual income than those in West Auckland and other regions when the level of education is high.
- 5. As a contrast, South Aucklanders who work in the Education/Training and Health/SocialCare sectors seem to have a lower average annual income than those in the other regions. Such difference is more significant when the level of education is high.

 Storyline 2:
 For local boards in South Auckland and West Auckland, does an industry, occupation or role have a higher (or lower) average earning per hour when compared to the rest of Auckland?



*Figure 3 Distribution of hourly income vs years of education in Auckalnd* 

### Key insights:

- 1. For the same industry and same occupation, South Aucklanders generally have a lower education length compare to that of West Auckland and the rest of Auckland.
- 2. South Aucklanders and West Aucklanders have higher hourly rate than the rest of Auckland in Manufacturing but seem to earn less per hour in Retail Trade, Professional/Scientific/Technology, Education/Training and Health/Social Care.
- 3. Rest of Auckland has a higher proportion of people who worked in the Health/Social Care industry than South and West Auckland.

South Auckland Top 5 roles that have less hourly rate than the rest of Auckland for each region								
Role, i.e. industry/occupation	Mean	Mean no.	Working	Mean	Mean no.	Working	The	
	hourly	Of working	pop size	hourly rate	Of working	pop size in	difference	
	rate	hours	within	in the rest	hours	the rest of	in hourly	
	within	within	South	of Auckland	within the	Auckland	rate (\$/hr)	
	South	South	Auckland		rest of			
	Auckland	Auckland			Auckland			
Agriculture/Forestry/Fish +	\$19.97	39.62	27	\$24.42	41.44	183	\$4.45(-)	
Clerical & Administrative								
Electricity/Gas/Water/Waste +	\$25.07	41.47	30	\$28.39	41.52	699	\$3.32(-)	
Professionals								
Transport/Postal/Warehouse +	\$23.81	42.00	240	\$26.55	40.71	69	\$2.74(-)	
Community & Personal Service								
Public Admin/Safety + Labourer	\$22.31	42.91	54	\$24.82	43.16	24	\$2.51(-)	
Education/Training + Technician	\$20.15	39.48	114	\$22.47	41.03	498	\$2.32(-)	
& Trades								
	1	1						

South Auckland Top 5 roles that have more hourly rate than the rest of Auckland for each region							
Role, i.e. industry/occupation	Mean hourly rate within South Auckland	Mean no. Of working hours within South Auckland	Working pop size within South Auckland	Mean hourly rate in the rest of Auckland	Mean no. Of working hours within the rest of Auckland	Working pop size in the rest of Auckland	The difference in hourly rate (\$/hr)
Accommodation/Food + Machinery Operation & Driver	\$21.05	40.44	66	\$18.07	41.21	63	\$2.98(+)
Information/Media/Telecom + Labourer	\$21.96	39.88	27	\$19.72	39.02	42	\$2.24(+)
Electricity/Gas/Water/Waste + Labourer	\$22.08	44.58	45	\$20.76	44.95	66	\$1.32(+)
Art/Recreation + Professionals	\$26.81	40.30	63	\$25.63	43.17	486	\$1.18(+)
Information/Media/Telcom + Machinery Operation & Driver	\$21.90	41.14	21	\$20.75	40.13	39	\$1.15(+)

West Auckland Top 5 roles that I	West Auckland Top 5 roles that have less hourly rate than the rest of Auckland for each region							
Role, i.e. industry/occupation	Mean hourly rate within West Auckland	Mean no. Of working hours within West	Working pop size within West Auckland	Mean hourly rate in the rest of Auckland	Mean no. Of working hours within the rest of	Working pop size in the rest of Auckland	The difference in hourly rate (\$/hr)	
Wholesale Trade + Community & Personal Service	\$19.79	42.95	21	\$22.30	42.08	87	\$2.51(-)	
Agriculture/Forestry/Fish + Clerical & Administrative	\$22.19	40.19	27	\$24.42	41.44	114	\$2.23(-)	
Finance/Insurance + Community & Personal Service	\$21.91	43.88	33	\$24.01	40.83	147	\$2.10(-)	
Rental/Hiring/Real Estate + Sales	\$20.55	43.66	144	\$22.55	43.17	750	\$2.00(-)	
Electrocity/Gas/Water/Waste + Sales	\$20.53	43.93	30	\$22.47	42.55	87	\$1.94(-)	

West Auckland Top 5 roles that have more hourly rate than the rest of Auckland for each region								
Role, i.e. industry/occupation	Mean hourly rate within West Auckland	Mean no. Of working hours within West Auckland	Working pop size within West Auckland	Mean hourly rate in the rest of Auckland	Mean no. Of working hours within the rest of Auckland	Working pop size in the rest of Auckland	The difference in hourly rate (\$/hr)	
Accomodation/Food + Professionals	\$26.88	42.09	78	\$24.64	43.02	282	\$2.24(+)	
Art/Recreation + Professionals	\$27.71	41.26	132	\$25.63	43.17	483	\$2.08(+)	
Public Admin/Safety + Labourer	\$26.83	43.18	27	\$24.82	43.16	84	\$2.01(+)	
Information/Media/Telcom + Clerical & Administrative	\$25.54	40.93	201	\$23.62	41.24	768	\$1.92(+)	
Retail Trade + Community & Personal Service	\$20.71	40.42	96	\$19.28	41.60	300	\$1.43(+)	

### Key insights:

- Comparing .to the rest of Auckland, South Aucklanders have the most disadvantaged in the industry of Agriculture/Forestry/Fish + Clerical & Administrative, and on average South Aucklanders less earn \$4.45 per hour. Moreover, in those industries that South Auckland people earn better than the rest of Aukland, South Aucklanders perform the best in Accommodation/Food + Machinery Operation & Driver, however, they merely more earn \$2.98 per hour more than the rest of Auckland people.
- 2. The difference between West Auckland and the rest of Auckland is relatively smaller than the difference between South Auckland and the rest of Auckland. In general, West Aucklanders perform worse in the sales and service industries and perform better in professional and administrative sectors.

# **Storyline 3:** What are the top 5 roles that TSI should focus their efforts on to encourage the potential workforce to get into?

Consider a ranking based on assigning equal weights to population size (i.e. the larger, the better), education level (i.e. the lower the investment in effort the better), and average annual income (i.e. the higher the better); we have identified the following top 5 occupation/industry combinations for South Auckland.

Occupation/Industr y	Working population size within the region	The average earning p.a. within the region	The average level of education within the	Ranking (based on average quantile score)
Construction + Machinery Operation & Driver	411	\$51,900	1.5	81.5
Transport/Postal/W arehouse + Machinery Operation & Driver	1,824	\$49,200	1.8	78.5
Manufacturing + Machinery Operation & Driver	2,001	\$48,400	1.9	75.8
Manufacturing + Clerical & Administrative	1,002	\$51,600	2.5	74.0
Manufacturing + Technician & Trades	1,998	50,600	2.6	71.3

Using a similar method, we have also identified the following top 5 occupation/industry combinations in West Auckland region.

Occupation/Industr	Working pop	The average	The average	Ranking (based
у	size within the	earning p.a.	level of	on average
	region	within the	education	quantile score)
		region	within the	
			region	
Manufacturing +	633	\$57,500	3.3	70.0
Manager				
Construction +	249	\$52,800	1.5	69.0
Machinery				
<b>Operation &amp; Driver</b>				
Manufacturing +	1,752	\$50,900	2.9	69.0
Technician & Trades				
Construction +	1,248	\$52,000	3.0	67.8
Technician & Trades				
Transport/Postal/W	606	\$48,800	2.0	67.3
arehouse +				
Machinery				
<b>Operation &amp; Driver</b>				

We understand that the Auckland Council TSI team has made some 'future growth' projections of different industries, mainly based on (future) population size increases. We aim to work with TSI subject matter experts (SMEs) on integrating this with that part of the analysis.





South Auckland Top 5 Industry/Occupation Standardised Age Distribution

Figure 4 South Auckland Top 5 Industry/Occupation Age Distribution. The distribution has been standardised by dividing the population of each age-group of the top 5 industry/occpation by the total population of South Auckland per age-group.



South Auckland Top 5 Industry/Occupation Standardised Ethnicity Distribution

Figure 5 South Auckland Top 5 Industry/Occupation Ethnicity Distribution. The distribution has been standardised by dividing the population of each ethnic-group of the top 5 industry/occpation by the total population of South Auckland per ethnic-group.



South Auckland Top 5 Industry/Occupation Income Distribution

Figure 6 Income Distribution of Top 5 Industry/Occupation in South Auckland



Geo-heatmap of population size living within each <u>meshblock</u> of South Auckland that worked in the top 5 industries/occupations

Figure 7 Geo-heatmap of population size per meshblock in South Auckland

### South Auckland region:

- 1. 16.5% of the *45-54 years old* and of the *55 64 years old* working populations within the South Auckland region are working on those 'top 5' industries identified above. (It seems South Auckland has a older working population in general).
- 2. Over 18% of the Pacific Island working population resides in South Auckland and work on those 'top 5' roles identified, followed by Maori (15.5%) and Asian (12%).
- 3. The average annual income of those who worked on the 'top 5' roles in South Auckland is ~48.5K p.a.

0.14



West Auckland Top 5 Industry/Occupation Standardised Age Distribution

Figure 8 Standardised Age Distribution of South Auckland Top 5 Industries/Occupations West Auckland Top 5 Industry/Occupation Standardised Ethnicity Distribution



Figure 9 Standardised Ethnicity Distribution of West Auckland Top 5 Industries/Occupations



Figure 10 Income distribution of West Acuklanders who worked in the top 5 industries/occupations



Geo-heatmap of population size living within each <u>meshblock</u> of West Auckland that worked in the top 5 industries/occupations

Figure 11 Geo-heatmap of population size within each meshblock of West Auckland

### West Auckland region:

- 1. 12.5% of the 65+ years old and 10.5% of the 15 24 years old working populations within the West Auckland region are working on those 'top 5' industries identified above.
- 2. Close to 11.5% of the Maori working population resides in West Auckland and work on the 'top 5' industries in the region. In comparison, only 6.5% of Asian working population are in the region.
- The average annual income of those who worked on the 'top 5' roles in West Auckland is ~50.2K p.a.

### Pilot Case Study:

To further investigate the equity of income, we have chosen the people who work as professionals in the Health/SocialCare industry and live in South Auckland. Their annual gross income is compared by gender and then by ethnicity.

A statistical test<sup>3</sup> was conducted to examine the income difference between males and females, and a p-value<sup>4</sup> of less than 0.05 was obtained. This suggests that the difference is significant and females earn at least \$11,368 less than males per year. The income differences by ethnicity were checked using an ANOVA test<sup>5</sup>, and the results showed that there was no significant difference in income between ethnic groups in this industry/occupation within South Auckland.

### **Income Sources**

Both the Census2013 and IR data sets record the income source type per person. We have chosen to use the grouping based on the IR records as it is more complete and less prone to human error. Also, the census is filled in by individuals and may not contain the most accurate data due to societal consideration (e.g. one might be less likely to honestly answer whether he or she is on benefit while earning wages).

Initially, we focused on the number of income sources per individual, e.g. the IR data was processed to the format illustrated below. Each of columns 2 to 7 represents an income source and takes either a 'Yes' or 'No' response.

UID	Has W&S	Has BEN	Has PEN	Has WHP	Has STU	Has PPL	Has CLM
X0001	Yes	Yes	Yes	No	No	No	Yes
X0002	Yes	Yes	No	No	Yes	No	No
		•••	•••	•••	•••	•••	•••

However, this format has the following limitations:

- The total number of all possible income source combinations is enormous. For instance, the total number of two-income source combinations is already 7 \* 6 = 42. Without TSI's guidance on which combinations to focus on, this presents a challenge for finding the most interesting ones for a further deep dive.
- 2) The number of income sources does not reflect the proportion of each income source. For example, assuming there are two individuals that both have W&S (wage and salary) and BEN (benefits) as their income sources. Person A has 90% W&S and 10% BEN; while person B has 5% W&S and 95% BEN. Then these two people shall be considered as different as their 'main' income source is somewhat different.

<sup>&</sup>lt;sup>3</sup> To be precise, a two-sample t-test was conducted here. The test is a formal testing procedure to test if there is any statistical difference in (mean) income level between two disjoint groups (e.g. female and male).

<sup>&</sup>lt;sup>4</sup> The outcome of a two-sample t-test is captured by p-value, a probability signifies how 'significant' the difference (between two groups) is. A general rule-of-thumb is any p-value < 0.05 (i.e. 5%) is considered as significant difference being observed.

<sup>&</sup>lt;sup>5</sup> This is another statistical test similar to two-sample t-test, but to be used when there are more than two groups (e.g. 3+ ethnici groups in this case).

Hence analysing the proportion of income from the different sources for the individuals would provide us with more meaningful insights. The following data set, which shows the income source composition by the individual, was generated to support this analysis.

UID	Percent of W&S	Percent of BEN	Percent of PEN	Percent of WHP	Percent of STU	Percent of PPL	Percent of CLM
X0001	90	5	4	0	0	0	1
X0002	10	60	0	0	30	0	0

We then applied a cluster analysis method to find the 'best' groupings of individuals based on their income composition. Such grouping (or cluster membership) would be one of the key attributes in our analysis below.

# **Storyline 1:** In South Auckland and West Auckland, what is the proportion of the working population that have 1+ sources of income? Amongst these multiple sources of income, what are the most common in terms of population size?



#### Figure 12 Income Sources Pie Chart of South Auckland

West Auckland



Figure 13 Income Sources Pie Chart of West Auckland

	Income	Proportio		Income	Proportio
	source(s)	n within		source(s)	n within
		the			the
		region			region
South	W/S	0.54	West	W/S	0.57
Auckland		0.071	Auckland		0.054
	VV/S +	0.071		VV/S+	0.054
	BEN			BEN	
	W/S +	0.024		W/S +	0.028
	PEN			PEN	
	W/S +	0.013		W/S +	0.012
	STU			STU	

### Key insights:

- 1. 86.82% of the South Auckland working population have a single income source. Within which, the most common income source is Wage and Salary (54%).
- 2. The most common 2-source income combination for South Auckland is Wage & Salary + Benefits (7%), followed by Wage & Salary + Pension (2%).
- 3. Similar patterns exist for the West Auckland region, e.g. ~86.97% of the working population have a single source of income, 57% of that has come from Wage & Salary only.
- 4. The most common 2-source income combination for West Auckland is Wage & Salary + Benefits (5.4%), followed by Wage & Salary + Pension (2.8%).



Figure 14 Boxplot of Annual Gross Income vs Income Sources per employment status of South Auckland



Figure 15 Boxplot of Annual Gross Income vs Income Sources per employment status of West Auckland



Figure 16 Boxplot of Annual Gross Income vs Income Sources per employment status of the Rest of Auckland

South Auckland (Total Population Count = 175473)									
	BEN	CLM	PEN	PPL	STU	W&S	WHP		
Full-time	4,029	1,821	2,706	1,131	549	67,929	2,667		
Part-time	3,195	303	1,704	222	1,065	13,962	720		
Unemployed	7,902	132	165	30	1,392	6,945	711		
Not in labour         21,087         663         16,158         471         2,769         13,956         1,089									

West Auckland (Total Population Count = 156711)								
	BEN	CLM	PEN	PPL	STU	W&S	WHP	
Full-time	2,880	1,644	2,856	1,248	456	64,887	3,672	
Part-time	2,973	315	2,019	264	1,212	15,183	1,107	
Unemployed	4,836	78	141	15	888	4,752	624	
Not in labour force	13,527	585	16,752	561	2,238	9,936	1,062	

Rest of Auckland (Total Population Count = 612747)									
	BEN	CLM	PEN	PPL	STU	W&S	WHP		
Full-time	7,647	5,856	15,486	4,158	1,824	273,147	16,308		
Part-time	8,319	1,329	11,187	1,071	5,187	68,184	5,463		
Unemployed	11,052	264	555	39	2,748	15,936	1,824		
Not in labour force	Not in labour force         30,117         1,689         71,259         1,782         7,977         38,442         3,861								

Besides, the demographic characteristics of the part-time working population in Auckland are summarised by ethnicity and age group in the following:

Ethnicity	Proportion (part-time population/total population of Auckland)
EU	12%
Maori	8%
PacIsd	6.5%
Asian	9.9%
Melaa	9.3%

Age Group	Proportion (part-time population/total population of Auckland)
15 - 24	18.7%
25 -34	9.6%
35 - 44	13%
45 - 54	13%
55 - 64	13.6%
65+	9.8%

### Key insights:

- 1. There are some similarities across South, West and the rest of Auckland, such as the population of not in labour force has the highest annual income in the CLM and BEN groups.
- 2. Full-time workers earn the most in the W&S and WHP groups, and the income of South Auckland is lower than in other regions in Auckland.
- 3. PEN and PPL look similar across the three regions.

### **Storyline 3:** What are the key factors that contribute the most to the clustering of individuals?

We have performed a two-tier clustering analysis for individual income profile:

a. We first classify the total ~800K population into 4 clusters with the following income profile descriptions. (Note: The total working population size with IR records in 2013 = ~800K)

BEN (i.e. benefits)	PEN (i.e. pension)	WHP & STU	WS (wages/salary)
80,946	120,864	41,814	555,861

b. Within each cluster in a. above, we further perform a clustering analysis to separate (large) clusters into subclusters.

BEN Only	BEN & PEN	BEN & STU & PEN	PEN Only	PEN & WHP	WHP & PEN	WHP & STU	WS Only	WS & BEN	WS & PEN	WS & STU
67,431	1,029	2,169	113,148	570	14031	41,811	513,927	22,152	6,114	17,097

c. We can then associate demographic attributes such as age, ethnicity, industry and occupation etc with the resulting clusters, and identify variables that contribute the most to each cluster.

### Key insights:

- 1. This two-tier clustering solution has identified 11 clusters successfully.
- It seems the key contributing factors that separate the clusters are 'age', 'occupation', 'industry' and 'employment status' (i.e. full-time, part-time etc). Qualification is also a variable that helps separate the clusters, but it has a relatively small influence compared to age, occupation etc.
- 3. Local board (or region of Auckland) does not seem to be one of the critical variables given age, and ethnicity is included in the clustering model. I.e. All local boards (or regions) seem to have a similar cluster distribution. Furthermore, when excluding age and ethnicity from the clustering model, the local board still has a very insignificant influence on the 4 clusters (first tier), and no influence on the 11 clusters (second tier).

# **Storyline 4:** For each cluster of individuals with similar income profiles, what are the typical demographic characteristics for them; and where do they live?

### Cluster 'BEN Only', e.g. mostly benefit takers:



Cluster BEN Mainly Standardised Age Distribution

Figure 17 Standardised Age Distribution of Cluster BEN



Cluster BEN Mainly Standardised Ethnicity Distribution

Figure 18 Standardised Ethnicity Distribution of Cluster BEN



Figure 19 Income Distribution of Cluster BEN

Where the benefit recipients lives in Auckland



Figure 20 Geo-heatmap of population size in each meshblock of cluster BEN

- 1. 17% of the people in the 55-64 years old age group in Auckland are classified as 'mostly benefit takers', followed by 11.9% of 45-54 years old and 10.5% of 35-44 years old.
- 2. 22% of Melaa in Auckland are classified as 'mostly benefit takers', followed by 21% Maori and 19% of Pacific Island.
- 3. The income distribution indicates that most of the people in this cluster have an income of around \$15k per year.

29





**Cluster PEN Mainly Standardised Age Distribution** 





**Cluster PEN Mainly Standardised Ethnicity Distribution** 

Figure 22 Standardised Ethnicity Distribution of Cluster PEN

**Cluster PEN Mainly Income Distribution** 



Figure 23 Income Distribution of Cluster PEN

Where the pension recipients lives in Auckland



Figure 24 Geo-heatmap of population size in each meshblock of Cluster PEN

- 1. 85% of the 65+ years old people are classified as 'mostly pension takers' as expected. Besides, 5% of 55-64 years old and a tiny proportion of 45-54 years old are also classified into this cluster.
- 2. 22% European are classified as 'mostly pension takers', followed by 8% of Pacific Island and 7% of Asian.
- 3. The income distribution for this cluster indicates that most of the pension takers have an annual income of around \$16k ~ \$17k.

# The projection from 2013 to 2018

One of the significant limitations of this project is that the census 2018 data has not been released by Stats NZ yet, so the analysis and insights can only be based on census 2013 and IR 2013 data. To bridge the five years gap, we have brought IR 2018 into the analysis to estimate a temporal change between the two years, and to do so, and a few assumptions have to be made:

- 1. The residential address of the matchable population in Auckland stays the same from 2013 to 2018.
- 2. The industry and occupation of the matchable population in Auckland has not changed from 2013 to 2018.
- 3. The average income across a year is relatively stable since IRD has only released data for the first six months of 2018, the annual gross income has to be estimated based on these six months only.

Moreover, to minimise the possible noise of the data, we have only targeted Aucklanders that are aged between 25 to 55 and had full-time work in the year 2013. Based on the analysis, we found that the annual gross income for the targeted population has an average increase of \$6,333 for South Auckland, \$7,784 for West Auckland, and \$11,059 for the rest of Auckland. This finding shows that South/West Aucklanders have a relatively lower rate of increase compared to the rest of Auckland.

In addition to the difference of annual income by region, we also deep-dived into the top 5 industries/occupations of different regions that have a working population greater than 10 and also have the largest average difference between 2013 and 2018, the findings are summarised in the following tables:

South Auckland								
Industry	Occupation	Average difference in Annual Gross Income	Population					
Construction	Manager	\$21,999 (+)	237					
Rental/Hiring/RealEst	Sales	\$18,699 (+)	117					
Prof/Scientific/Technic	Professionals	\$17,301 (+)	699					
Construction	Professionals	\$16,800 (+)	132					
Agri/Forestry/Fish	Manager	\$15,801 (+)	24					

West Auckland							
Industry	Occupation	Average difference in Annual Gross Income	Population				
Rental/Hiring/RealEst	Sales	\$27,900 (+)	123				

Construction	Manager	\$23,901 (+)	387
Transport/Postal/War ehouse	Professionals	\$20,502 (+)	120
Construction	Professionals	\$20,301 (+)	201
Rental/Hiring/RealEst	Manager	\$18,099 (+)	147

Rest of Auckland				
Industry	Occupation	Average difference in Annual Gross Income	Population	
Rental/Hiring/RealEst	Sales	\$32,301 (+)	585	
Construction	Manager	\$27,798 (+)	1,893	
Rental/Hiring/RealEst	Manager	\$26,400 (+)	831	
Construction	Professionals	\$24,501 (+)	906	
Rental/Hiring/RealEst	Professionals	\$24,099 (+)	636	

Another interesting topic is the equity amongst ethnicities across all regions of Auckland, so we compared the average annual income difference across different ethnicities, and the findings are outlined below:

Ethnicity	Population	The average difference in Annual Income
EU	106,491	\$10,900 (+)
Maori	21,534	\$6,000(+)
Pacisland	20,301	\$5,200 (+)
Asian	48,690	\$10,900 (+)
Melaa	3,108	\$10,400 (+)
Other	3,237	\$9,500 (+)

We have to remind readers again that these figures are calculated based on a few assumptions, and they may not fully reflect the most current situations. The results could be improved when the census 2018 data and the full IR2018 data becomes available.

# **Dashboard Views**

In addition to the insights and analysis presented here, we have also created an interactive dashboard that allows users to explore data at their leisure. There are three main pages of the dashboard, namely:

- 1. Overview
- 2. Wage & Salary vs Education
- 3. Income Sources

The dashboard is built using the R/RShiny framework and is currently hosted by Harmonic. Once it is been deployed, it can be accessed through a web page (details are to be provided upon deployment).

The rest of this section captures some of the dashboard views and their key functionalities. The contents for this section will be updated and provided once the dashboard development is completed.



Southern Initiative	≡
<ul> <li>Executive Summary</li> <li>Comparative Analysis</li> <li>Income VS Education</li> <li>South Auckland</li> <li>West Auckland</li> <li>All Auckland</li> </ul>	Income VS Education by Industry (All Auckland)          100k       Image: Second the second th

### **Next Steps**

Several future project ideas have surfaced from this project:

- 1. Provide a hosting environment, on-going support, regular updates and enhancements to the dashboard.
- 2. Leverage other data sources to improve the analysis.
- 3. Further detailed analysis of specific areas of interests. E.g. breaking down benefit takers by different benefit types and carry out a similar analysis to what we have done here.
- 4. We have made some projections of income values based on 2013 data with crude assumptions (e.g. linear increase of incomes assuming working population do not change roles in between the five year period). If 2018 census data becomes available in the future, we can refresh the analytical results based on more recent data sets.