**Business Analytic Project Final Report**



**School of Information Technology**

**Year 2**

**IT2356 - Business Analytic Project**

**Prepared by:** Heng Li En (191081H)[Team Leader], Sim Kai Yi(201912Z), Javious Leong Yi Song(200241Q), Li Zheng(203774M)

**Module Group:** BA2001

**Lecturer: Ms Jacinta Ong**

Date of Submission: 18th August 2021

**Appendix F - Table of the Content**

**i) Executive Summary**

• Summary of business objectives, problem statement, hypothesis, or goals - PG 1

• Describe the business scenarios, recommendation and insights discovered - PG 2 - 6

**ii) Project Plan**

Project team organization - PG 7

Project schedule and task allocation (Gantt chart) -PG 7

Software tools used for development - PG 7

**iii) Data Understanding and Modelling (You may want revise Week 1’s Project slide again**)

• System Architecture Diagram (i.e. draw a diagram to show how data come from the sources and go through different stages of analysis) - PG 8

• Data Catalogue (i.e. similar to Appendix D) - PG 9

• Describe the analytics model(s) used (example, descriptive analytics, diagnostic analytics, predictive analytics, prescriptive analytics, etc.) and how the model results are linked to your hypothesis and make recommendations. - PG 10 - 13

• Dashboard (with instructions - and highlight any special features) - PG 14 - 20

**iv) Problems Encountered**

Problems Encountered - PG 21

**v) Future Enhancements**

Future Enhancements - PG 22

**vi) Conclusion**

• Summarize the works, provide results interpretation, and conclude whether the outcome of the analysis fit the project expectation - PG 23

## 

# **Executive Summary**

Our business objective was to discover insights on the social , economic impacts that are happening in Singapore, whether it affects the fertility rate in Singapore. Besides that , our team would derive our problem statement on what is causing a decline of birth rates. Possible factors include cost of living, childcare expenses, recessions, subsidies, cultural or work issues for us to prove or disprove our problem statement with the creation of hypotheses. Our goal is to deliver the information to our target audience which includes the public, government, so that Singaporeans are able to reference whether an individual is able to afford having a child.

Heng Li En Hypothesis:

1. The higher the cost of living, the lower the birth rates
2. The higher amount of child care centre would lead to lower prices , eventually lead to higher birth rates

Javious Leong Yi Song Hypothesis:

1. Recessions have an effect on birth rates
2. The job security of a nation affects birth rates.

Li Zheng Hypothesis:

1. Increasing the government’s subsidies and grants in short-term upfront expenses will increase the birth rates in Singapore
2. Increasing the government increase their subsidies and grants on long term expenses will increase the birth rates in Singapore

Sim Kai Yi Hypothesis:

1. Social attitude towards having a child has changed throughout the years, as in the recent decade it discourages people from having a child.
2. The More marriages each year and the younger someone is the more babies will be born the following year

### **Business Scenario**

Business scenarios are the public is uninformed and the inadequate support from the government.Therefore there is a need to educate and inform the public about it and let our findings be the call to action to the government.

This could also result in how an average Singaporean might struggle affording a child given the expenditure of cost of living is high. The decline of fertility rate is based on an individual decision or act.

### **Recommendations**

**Heng Li En Hypothesis:**

1. Transportation cost can be calculated to find the shortest and cheapest route for an individual to reach their location. With food and beverage as the second highest expenditure cost , there is a need to reduce it by advocating hawker centres or increasing the number of hawker centres to the general public.
2. By coming out with more childcare centres widely spread in Singapore, the chances of parents entrusting their child with childcare centres, and would potentially grow a sense of reliance from it.Domestic helper costs can also be reduced to cater families who might not be able to take care of their child 24/7. To add on, reduced baby cost necessities could also help parents lighten up their struggles with financial issues and also source out for cheaper options or alternatives to buy without a problem.

Insights Discovered:

1. Household Income is higher than expenses stating that the reason for not having a child is not because of the expenses incurred.
2. High cost of living still impacts the majority of Singaporeans' lifestyle when making a decision in having a baby.
3. The commitment level to take care of a child is mostly reliant on childcare centres and domestic helpers to reduce workload.
4. It is predicted that household expenses would be increasing , stating other expenditures would increase based on the current Singaporeans spending.
5. Looking in a real-world context that consumers demand has been increasing , thus more expenses incurred to cater to their needs.
6. Program Fees still take a toll on parents given how the prices have been spiking up over the years.
7. Target insights on their spendage trends were mostly on Transportation and Food and Beverage.

Javious Leong Yi Song Hypothesis:

1. The extent of a recession can affect the birth rates of the given time frame.
2. The job security of the nation in a given period affects the birth rates.

Javious Insights:

1. The monthly median income has been constantly increasing over the years.
2. GDP has also been constantly increasing over the years, in the same way as median income has increased.
3. There was a large spike in unemployment rate in 2008-2009 due to the 2008 recession.
4. None of the data has any strong relations to birth rates.
5. The job vacancy to unemployment ratio experienced a sharp drop in 2012 before reaching its peak in 2014.
6. The unemployment rate had a sharp increase going from 2008 to 2009.
7. The data has no relation to birth rates, only to each other.

Li Zheng Hypothesis:

1. To lighten the requirement that has to be met to be approved for the special CPF Housing Grant
2. To increase the government Medifund Grant amount on more costly medical expenses to reduce the fear of financial instability

Li Zheng Insights:

1. Price of HDB has remained stable while the Number of HDB been build is increasing, hence showing that there is a stable supply
2. The Total Approved special CPF Housing Grant has a very strong correlation with the fertility rate
3. Special CPF Housing Grant give the most amount at up to $80 thousand as applications eligible for it, is also given the Enhanced CPF grant alongside the Additional CPF Grant
4. The increase in total consumer spending is growing exponentially more as compared to the total admission to both public & private hospitals
5. Medifund Grant has been increasing from 2006 to 2014, after 2014 which was the peak, there was a sharp drop from $159.3 Million in 2014 to $143.9 Million in 2016. It took till 2019 to finally go back up to $159.1 Million
6. There is a weak correlation between the rise and fall of the Average Grant Amount Per person and the fertility rate

Sim Kai Yi Hypothesis:

1. The more marriages each year and the younger someone is the more babies will be born the following year
2. Social attitude In recent years can discourage people from having children

Sim Kai Yi Insights:

1. More marriages does not necessary translate into a higher fertility rate the following year
2. Different ethnicities tend to have different amounts of fertility rate accordingly.
3. Malay’s tend to have higher fertility rate despite having a decreasing marriage rate
4. Rest of the ethnicity have a consistent or increasing marriage rate but have a decreasing birth rate
5. Malay’s tends to have a bigger household size compared to the other races
6. People that are older tend to give birth to more children compared to younger generations
7. Younger people do not have a child compared to people that are older
8. When Flat prices decrease fertility tends to increase a bit the following years
9. Giving more leaves does not necessary increase the fertility rate
10. The more people working in the society, the higher the sucide rate for ages 20-29.
11. Singaporeans find that there is not enough time to help their children in their studies
12. By having lower working hours fertility rate tends to be higher to begin with
13. Singapore has the highest average working hours per week compared to other countries.

# 

# 

# **Project Plan**

**Project Team organization**

Heng Li En (Leader), Li Zheng, Javious Leong, Sim Kai Yi

**Project Schedule and Task Allocation (Gantt chart)**

**Legend:**

* Everyone (E)
* Kai Yi(KY)
* Li Zheng(LZ)
* Li En(LE)
* Javious (J)
* Individual Work(IW)

PJ1. Identify problem statement(E) & PJ2. Business Understanding Hypothesis(E) - Week 2

PJ2. Business Understanding Hypothesis(E) PJ3. Data Understanding(E) - Week 3

PJ3. Data Understanding & PJ4. Finalise the Target Problem and Hypothesis(E) - Week 4

PJ4. Finalise the Target Problem and Hypothesis(E) - Week 5

PJ5. Prepare Project Proposal(E) & PJ7-10. Data Preparation(IW) - Week 6

PJ6. Proposal Presentation(E) & PJ7-10. Data Preparation(IW) - Week 7

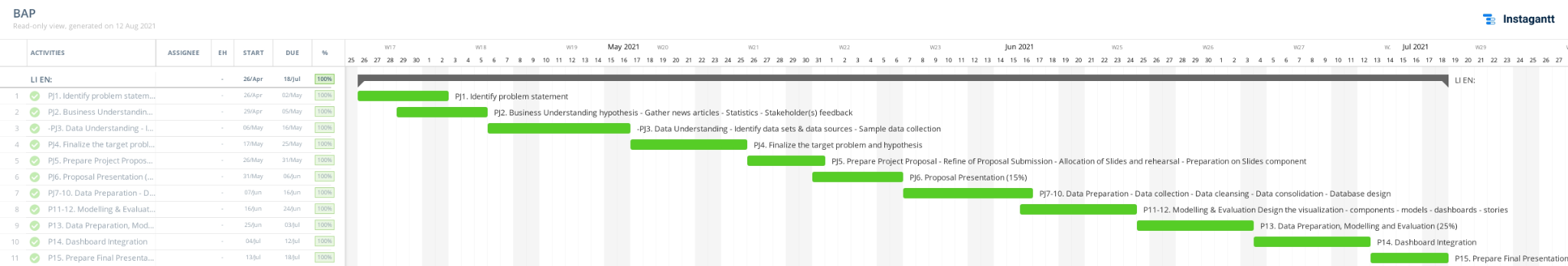
PJ7-10. Data Preparation (IW) - Week 8

PJ7-10. Data Preparation(IW) & work on Appendix D(E) - Week 9 & 10

PJ11-12. Modelling and Evaluation(IW) and Evaluation by teacher(IW) - Week 11-15

P13. Data Preparation, Modelling and Evaluation(E) - Week 16

P14. Dashboard Integration (E) & P15. Prepare Final Presentation,Poster,Story (E)- Week 17

****

**Software tools used for development**

Tools used are Excel, Tableau Prep Builder, Tableau and Python:

Excel and Tableau Prep Builder - Used to Clean data

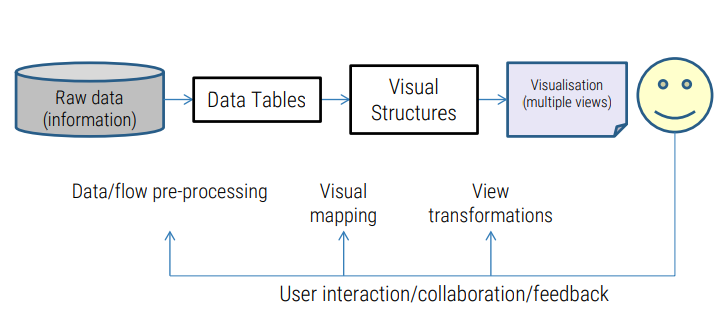
Tableau - For Data Visualisation and data modelling

Tableau - Used for Predictive Modelling (Li En)

Python - Used for Predictive regression(Li Zheng)

Excel - Used for Predictive regression (Kai Yi)

# **System Architecture Diagram**

****

Raw Data - Collection of Data from sites such as:

1. Data.gov.sg
2. Singstats.gov.sg
3. Statista
4. Samaritans of Singapore
5. Housing Board Development

Data/flow pre-processing between Raw Data and Data Tables includes Manual Download of data from stated sites and the usage of Tableau Prep to clean dirty data and constructing new data

Data Tables - Data tables in their respective excel workbooks

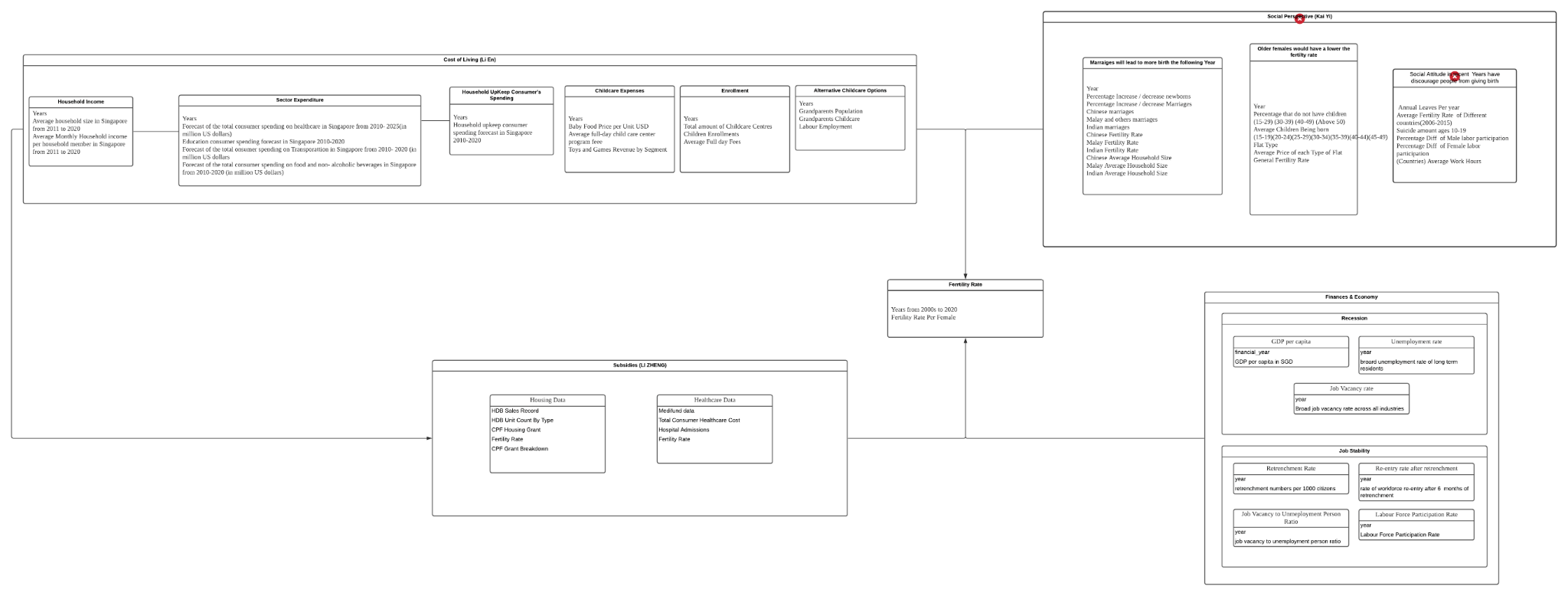
Visual Mapping - Testing of different visualisation or chart types and finding that best suit the data in helping it tell its story/insights

Visual Structure - visualizations of the previously cleaned data

View Transformation - Change the order of the visuals to help guide the flow of the story

Visualisation - Combining Multiple Visual Structures/Charts into a single dashboard to tell a story

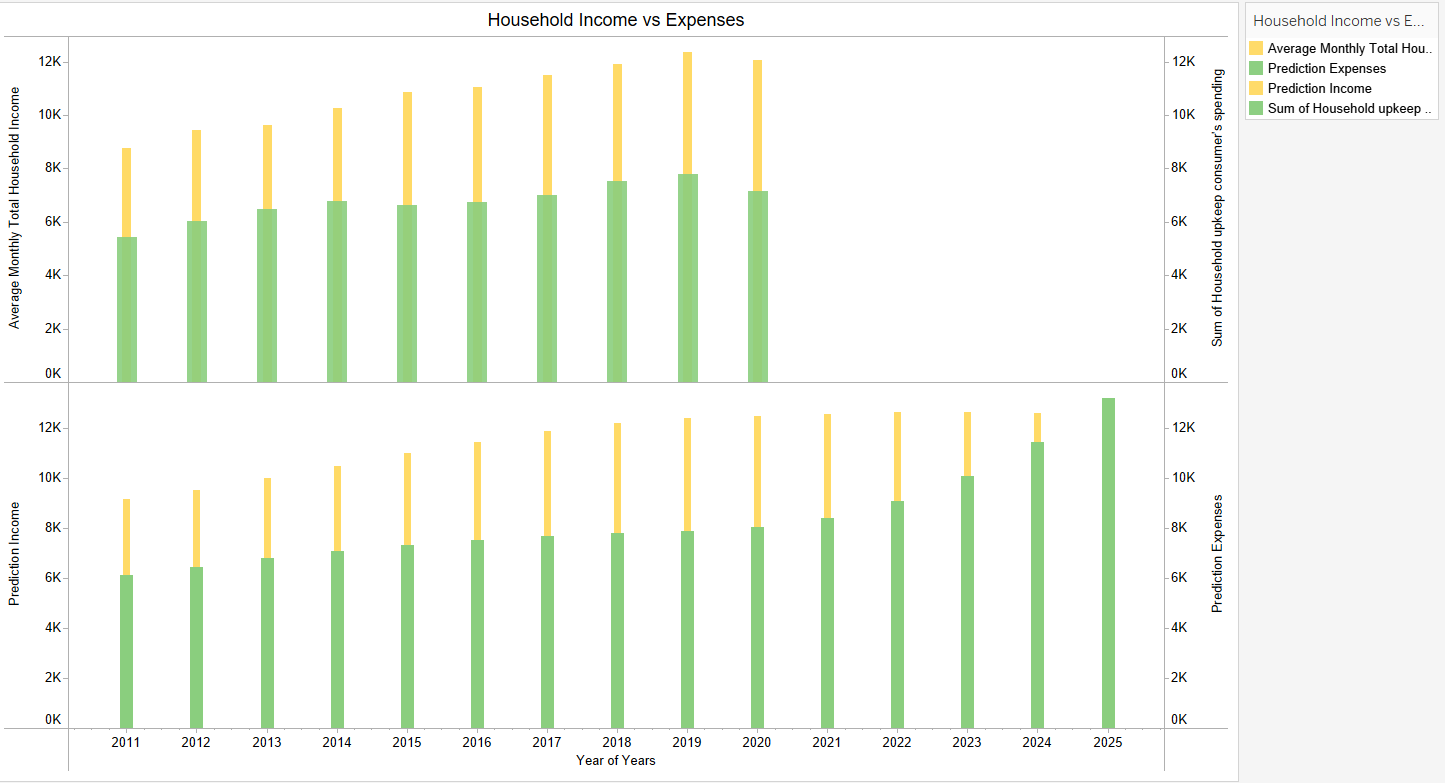
# **Data Catalog**

****

# **Analytics Model**

### Heng Li En

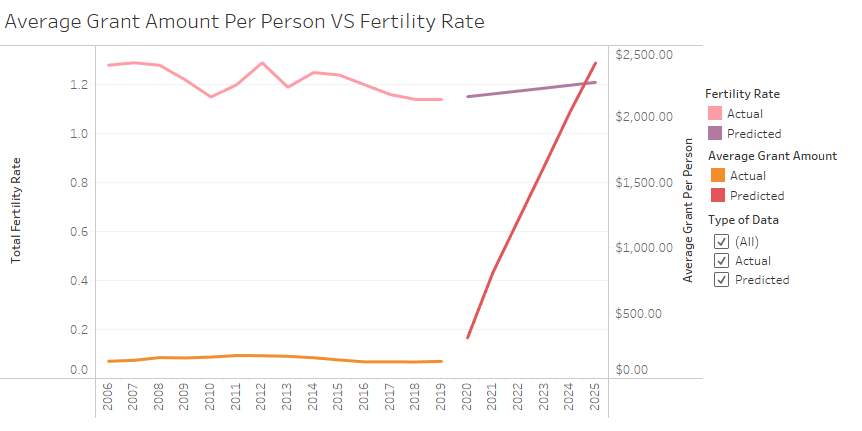
Predictive Analysis used in Household Income Vs Expenses Chart



I used predictive modelling to predict whether the expenses would decrease afterwards and continuously increase. It is proven that expenses will be increasing over the next few years and substantially would play a role in affecting adults on deciding to have a child. Given that household expenses would increase over the years, it is less likely the fertility rate would increase.This leads to saying that even with more savings , the chances of them wanting to have a child is low.This explained that singaporeans major spendage on household has been increasing.Thus to reduce the cost in household which includes utilities, household consumption on food and beverage,transportation cost and amount paid for lodging.

### Li Zheng

Linear Regression (Predictive Analysis) Used in Average Grant Amount Per Person VS Fertility Rate



In order to prove that if the government further increases their Medifund Grant Amount provided for the applicants, there will be an increase in the Fertility Rate. For the Linear Regression, for the training data, I used the average grant amount and the total healthcare consumer expenditure from 2010 to 2019 to use as the predictors and the fertility rate to be the end result. For the testing data, I used the training data of 2018 and 2019 to find the rough accuracy of the model which ended up having a standard deviation of 3% for the end result. From there, I used the estimated total consumer spending data from 2020-2025 and targeted fertility rate (1% more than the previous year) as the predictor to find the minimal amount of grant required by the model to get the targeted fertility rate.

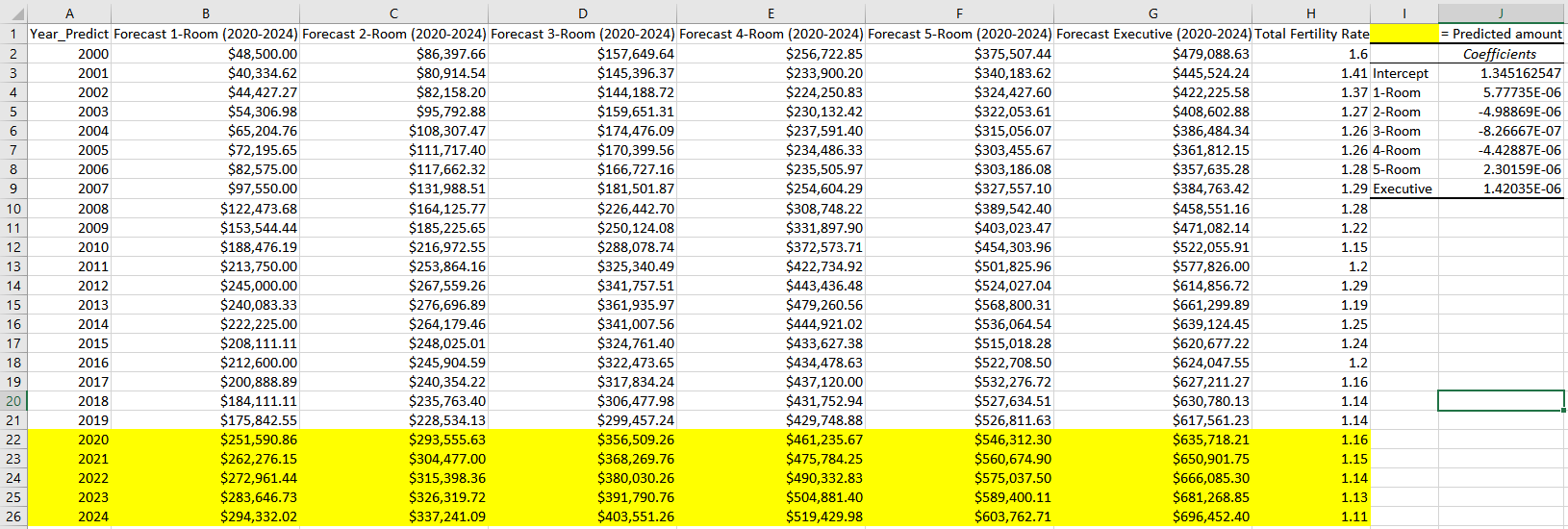
Code for Linear Regression:



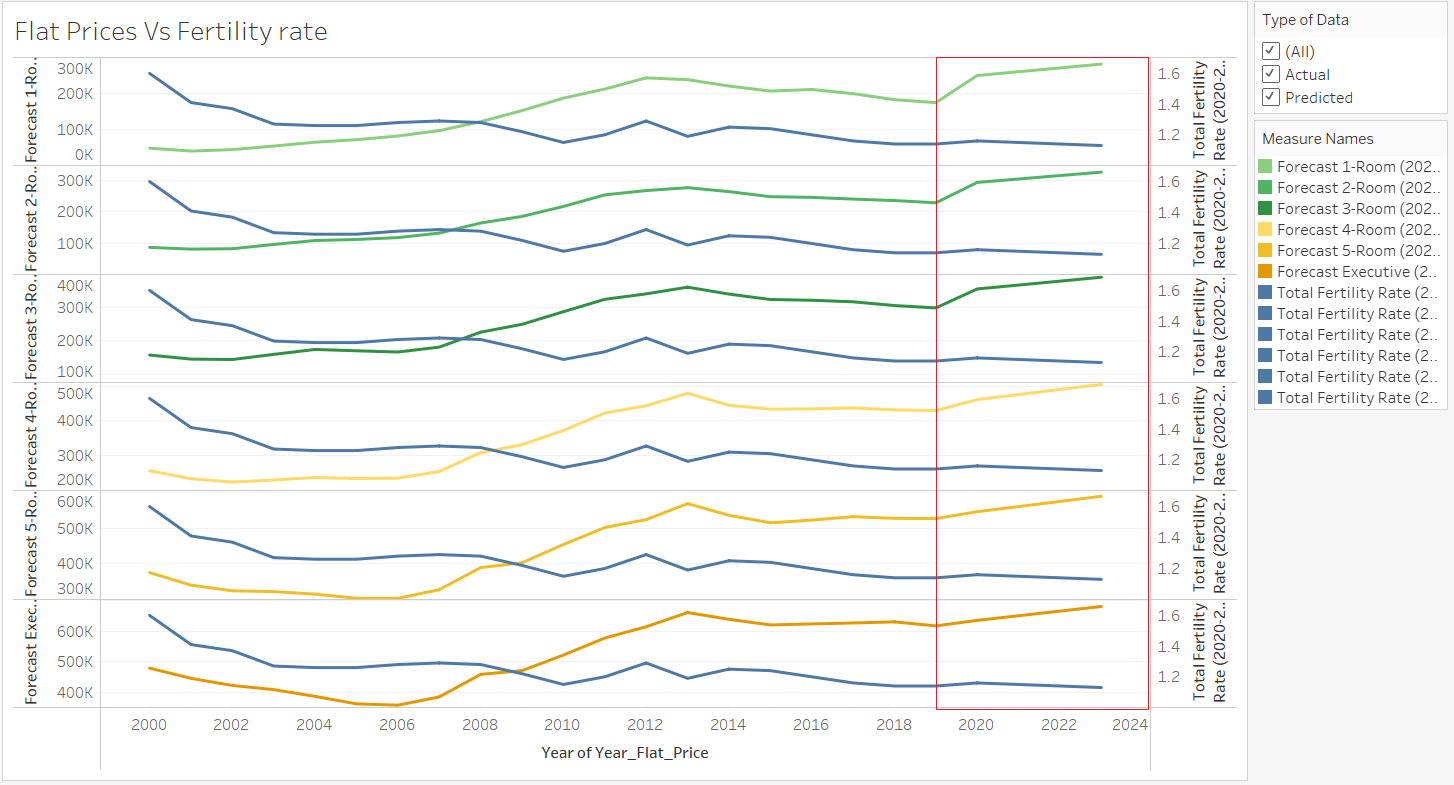
## 

## 

### Sim Kai Yi



In order to find the prices of future flat types I used the forecast function in excel to forecast the prices for 5 years (2020-2024). Then I used the forecast prices to find predictive regression using excel. By using linear regression, the formula *(predicted fertility rate = Intercept + 1-Room\* Forecast 1-Room + 2-Room\* Forecast 2-Room + 3-Room\* Forecast 3-Room + 4-Room\* Forecast 4-Room + 5-Room\* Forecast 5-Room + Executive \* Forecast Executive)* was used to find the fertility rate of 2020-2024.

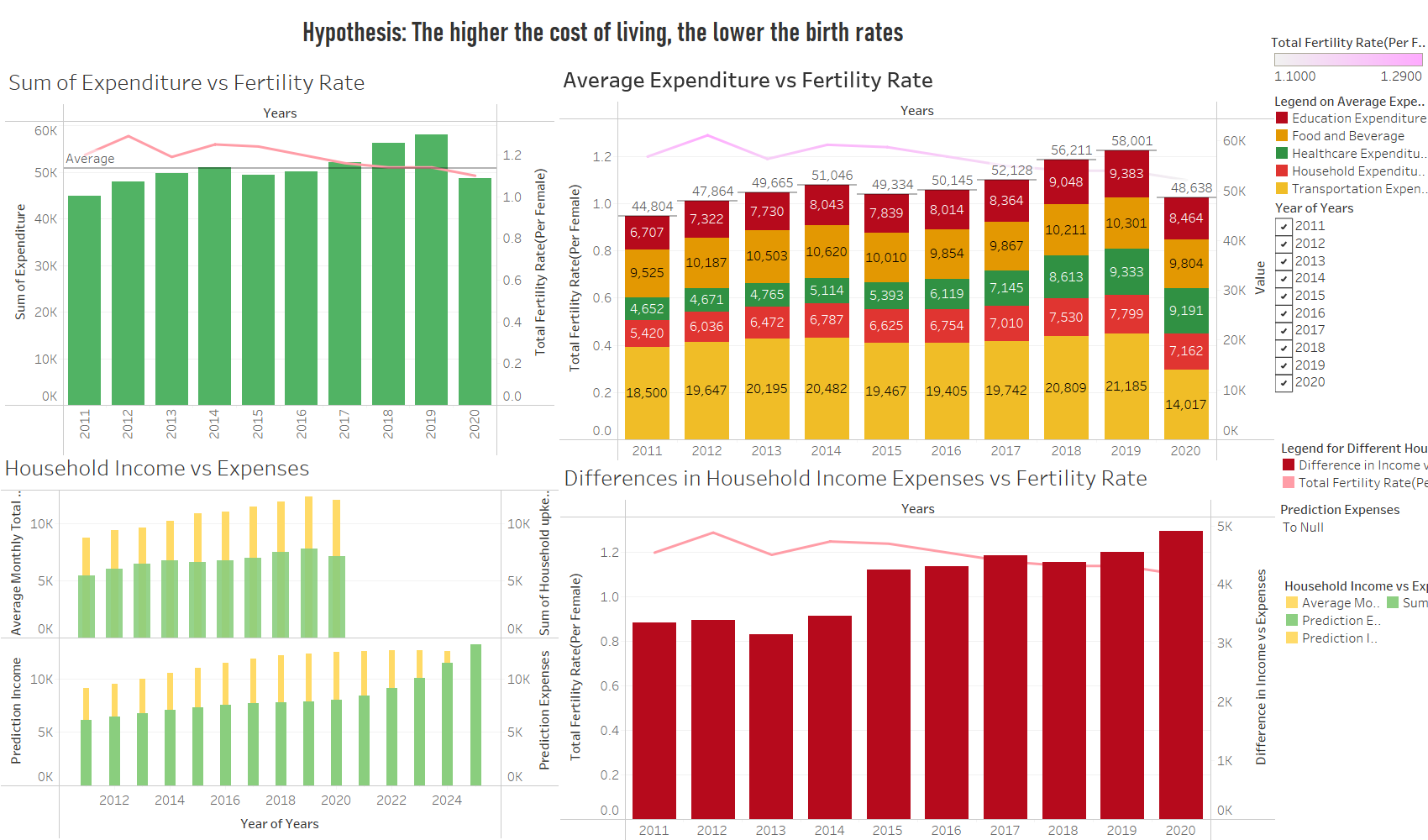


The red box labeled in the tableau visualisation shows the predicted fertility rate and forecasted flat prices. Then we can see clearly that when prices increase for different flat types the fertility rate would go down.

# **Dashboard**

### Heng Li En

Dashboard 1:

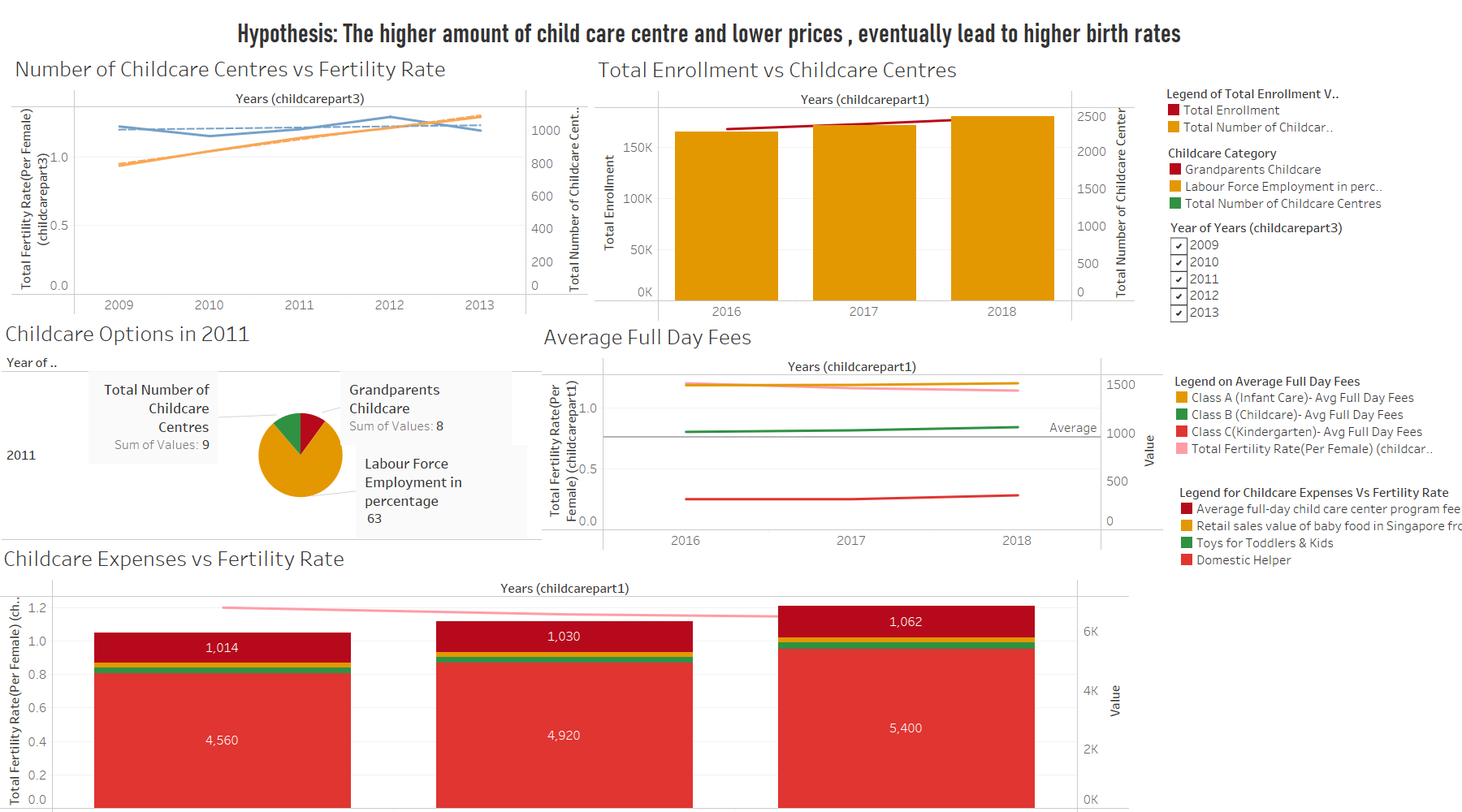


Special Features include:

* Use of legends and filters for clearer explanation
* Average Line in both Sum of Expenditure and Average expenditure vs fertility rate graph
* Predictive Modelling on Household Income Vs Expenses , so forecast 5 years later to see if the household expenses would exceed over the income and potentially affect the decision of the public.
* Ability to click charts and connect with one another for interactivity , Click to Filter

### Heng Li En

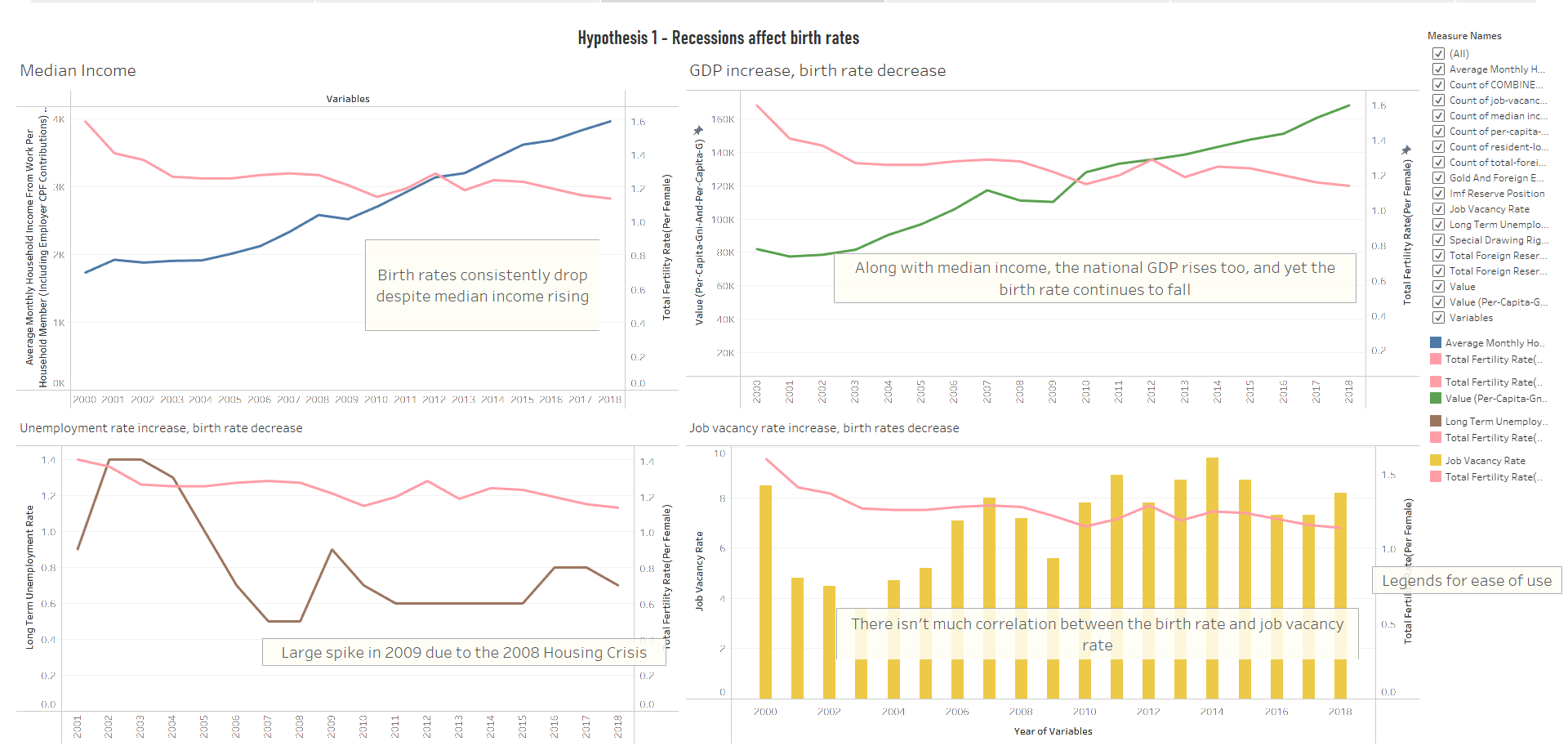
Dashboard 2:



Special Features include:

* Use of legends and filters
* Ability to click charts and connect with one another for interactivity
* Forecast on number of childcare centres vs fertility rate

### Javious Leong Yi Song

Dashboard 1:

Special Features include:

Legends

* Filters
* Clicking on a point connects to the other charts for increased interactivity
* Filtering in the entire dashboard

### Javious Leong Yi Song

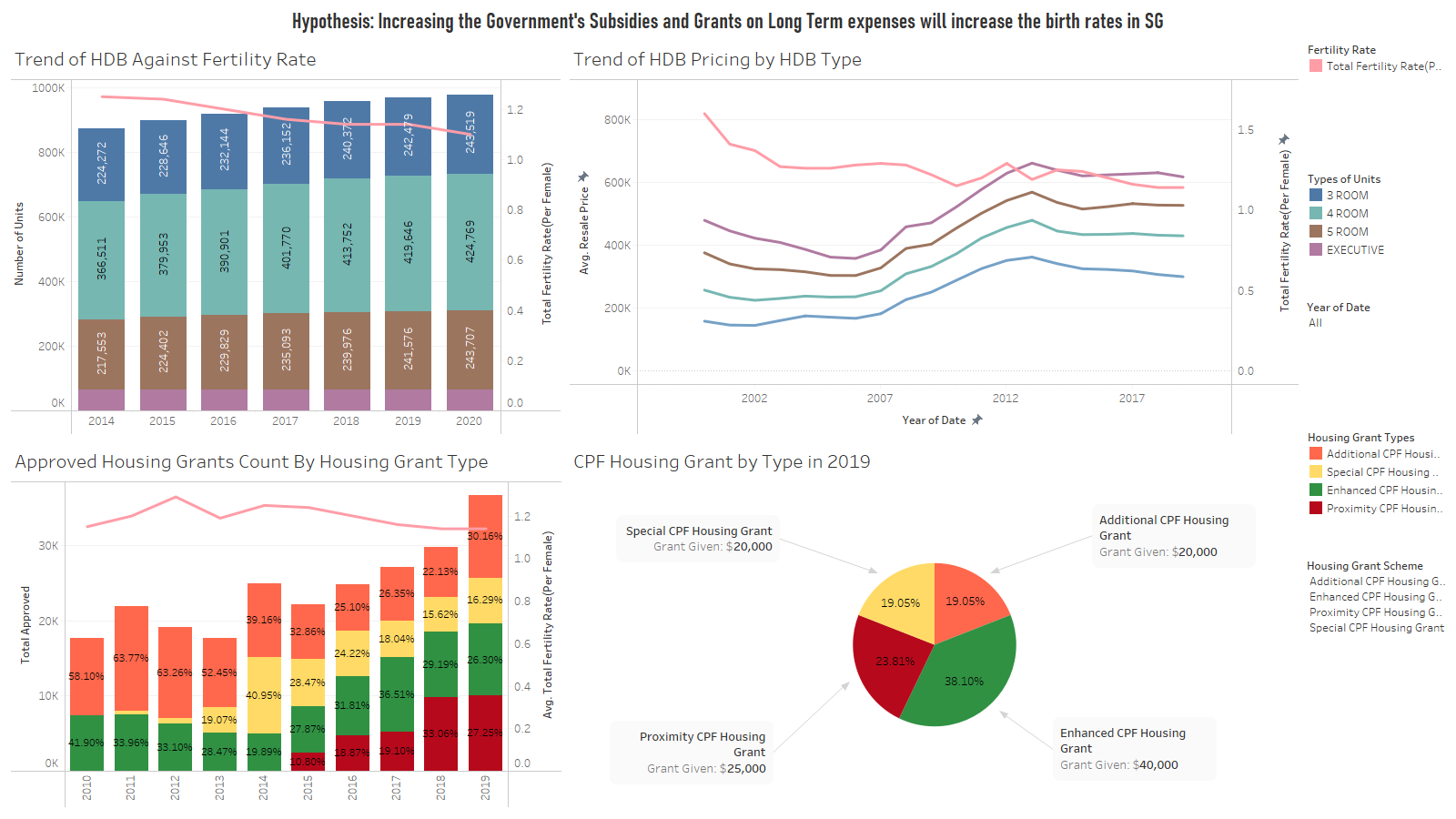
Dashboard 2:

Special Features include:

* Filters
* Legends
* Clicking on a point connects to the other charts for increased interactivity
* Filter for the entire dashboard

### Li Zheng

Dashboard 1

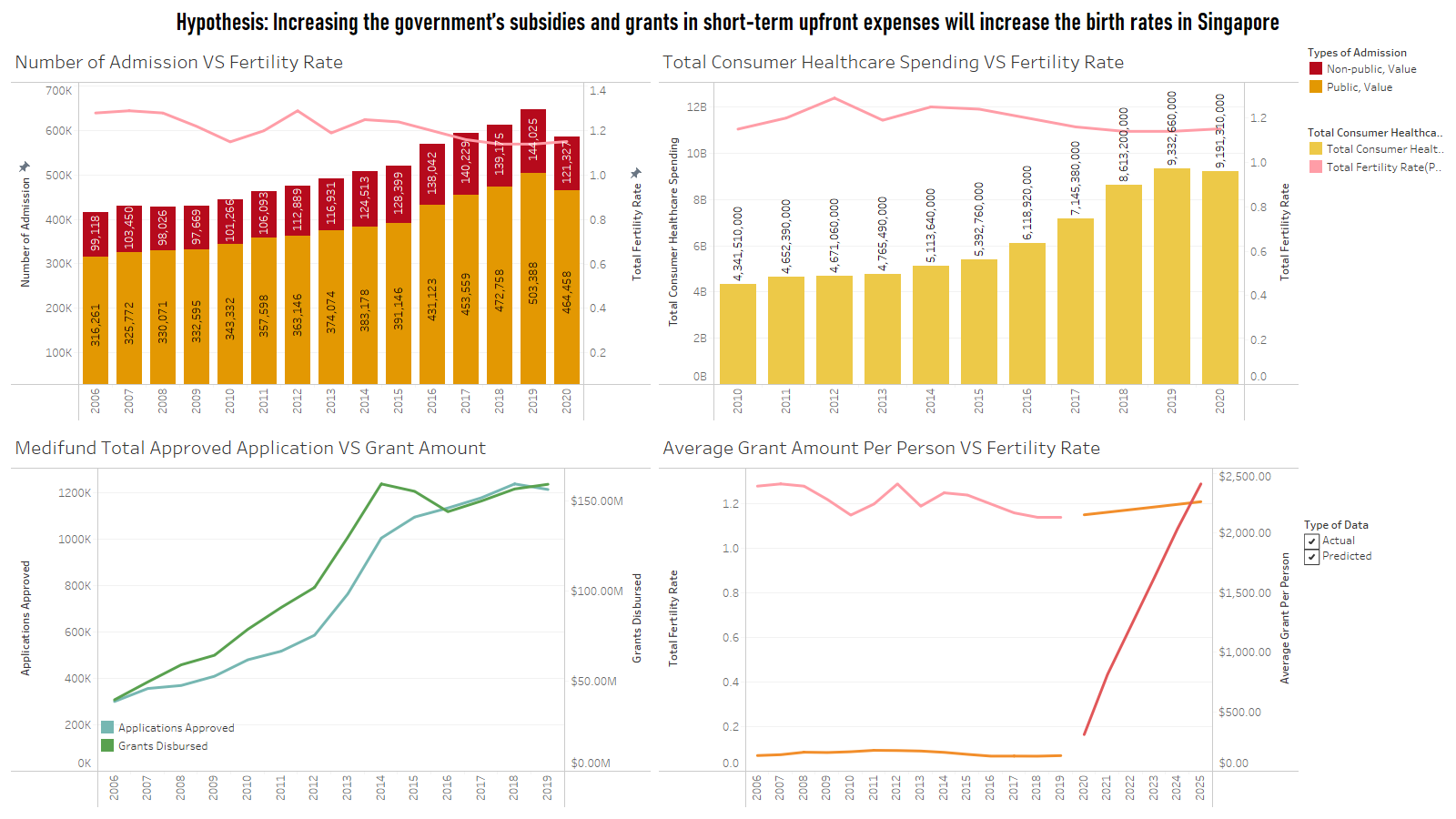


Special Features include:

* Use of legends
* Year Filter to drill into the different years
* Filter for the Housing grant Scheme to help filter the Housing Grant stacked bar chart
* Interactivity between the charts (Filter for entire dashboard)

### Li Zheng

Dashboard 2

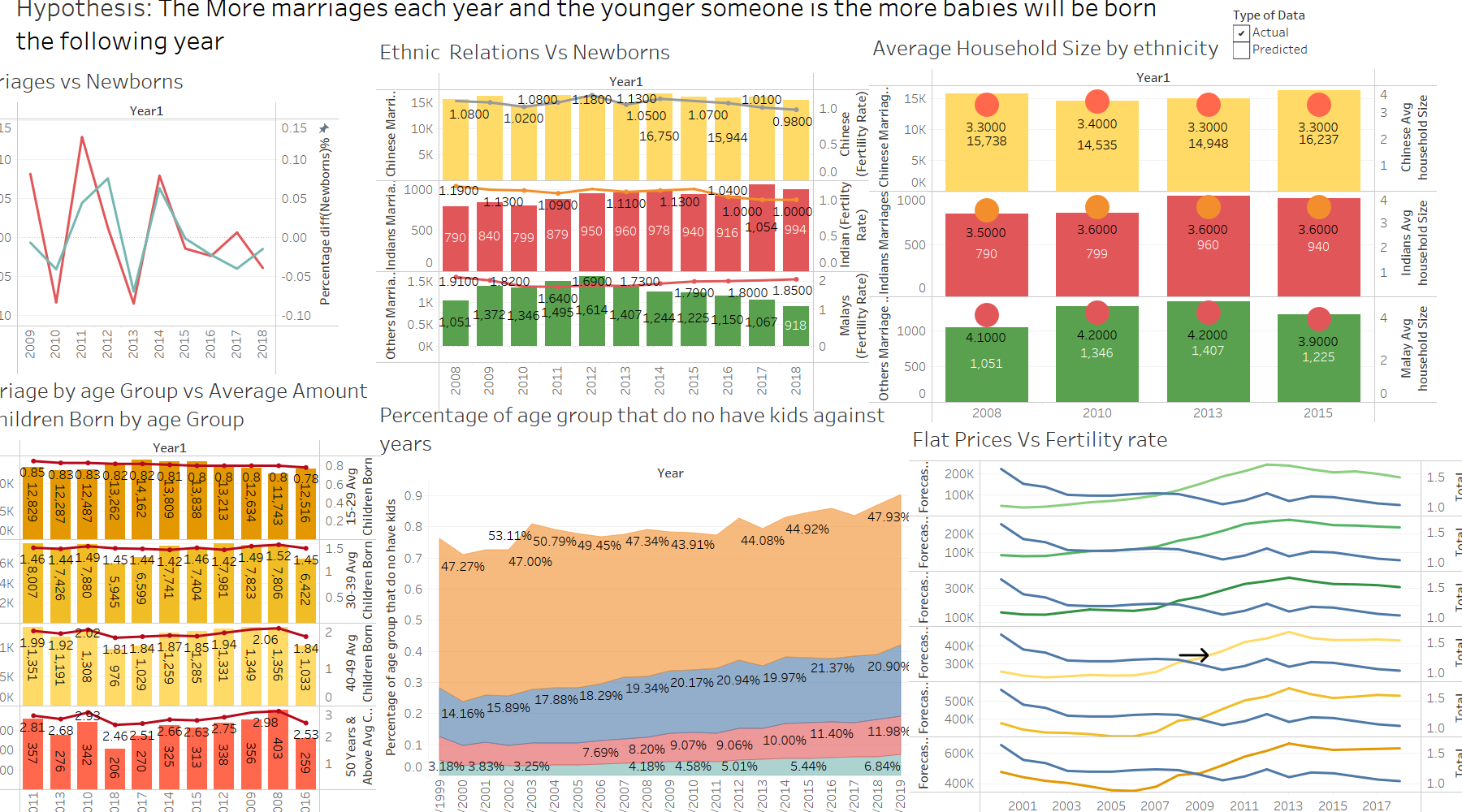


Special Features include:

* Use of legends
* Interactivity between the charts (Filter for entire dashboard)
* Type of data filters for the 4th chart (Average Grant Amount Per Person VS Fertility Rate) to separate and allow users to clearly see the actual data (collected data) and the Predicted Data (Use of Linear Regression Model to get the values)

### Sim Kai Yi

Dashboard 1:



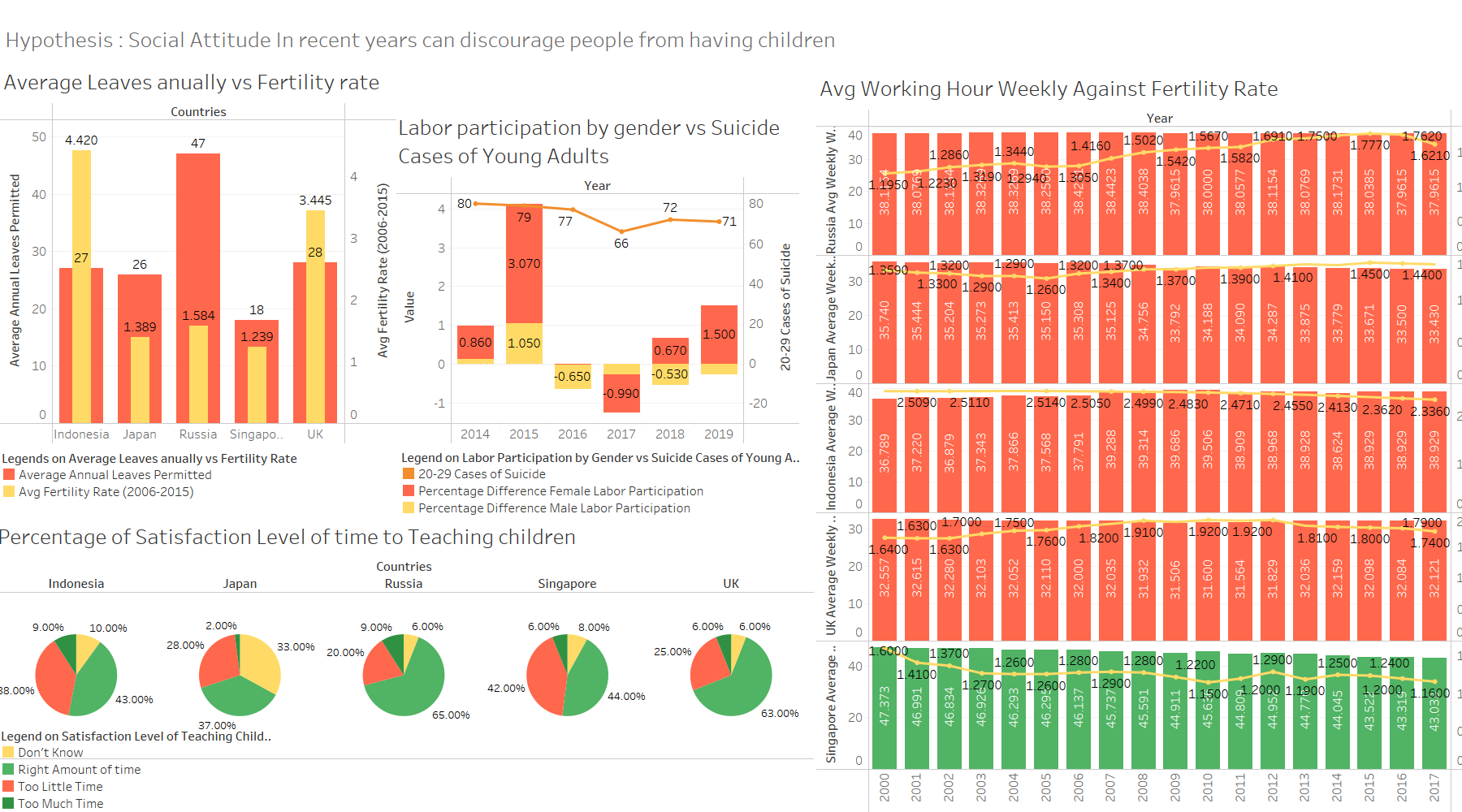
Special Features include:

* Global Filtering of other visualisations
* Predictive for flat prices and fertility rate from year 2020-2024
* Forecasting for prices of all flat types

### 

### Sim Kai Yi

Dashboard 2:



Special Features include:

* Measure Names
* Global Filtering of other visualisations when possible

## 

# **Problems Encountered**

These are a few points that we encountered when working on the story.

1. Integration and combining of tableau extracts did not work as there was a specific limit hence there is a need to consult the lecturer to ask if there are other methods to integrate them together.
2. Unable to do predictive modelling for certain charts and need to sought other ways of modelling which we have not learnt
3. The visualisations might not have enough space in a dashboard.So there is a need to rearrange or have floating charts.

1. Lack of data, therefore we can't work on predictive modelling or visualization charts we wanted to do initially as we are unable to source out the direct source from straits times, surveys hosted by private companies or the government.
2. We did not have enough user interaction or feedback.

# **Future Enhancements**

Some Enhancements we can make for carrying out our project:

1. To combat the lack of data, in the future, we should conduct surveys or user feedbacks for our topic in order to gather more data that we can use
2. Try out other visualisation softwares and use the one that offer more control over how much space each chart
3. In the future, to solve the file count limitation, we should aim to better combine multiple data table together to reduce the amount of joins between the different data tables
4. Take more time beforehand to learn and prepare ourselves for the use of the different types of predictive, descriptive and other forms of data mining methods

# Conclusion

In conclusion, there aren't enough grants or subsidies provided to an average Singaporean to increase the fertility rate given how the cost of living, cultural work environment or even recessions play a role in affecting it. This impacted Singaporeans decision in having a child given that the circumstances of having a child concluded more commitment, expenses to be incurred.

From the results,our team has derived that cost of living has impacted Singaporeans lifestyle in a certain way,and that high expenses on child necessities , with a small sum on medical grants, is not enough to cover Singaporeans spending per year. Because if recessions were to happen per year, the unemployment or retrenchment rate would be higher and affect Singaporeans in wanting to have a child. In addition , marriages also play an important factor whereby the lower the marriage rate , the lesser chance of fertility rate increasing.Thus the outcome of the analysis has matched up the project expectation which was that if we provide more financial support to the singaporeans, there would be a increase in fertility rate so there will be a reduced concern of financial instability.