



MONASH  
University

MONASH  
BUSINESS  
SCHOOL

Department of  
Econometrics &  
Business Statistics

☎ (03) 9905 2478  
✉ [BusEco-Econometrics@monash.edu](mailto:BusEco-Econometrics@monash.edu)

ABN: 12 377 614 012

# Sectoral Employment Dynamics in Australia

**(Elvis) Zhixiang Yang**  
EBS Honours Student

Report for  
Honours **Research Plan**, supervised by **Farshid Vahid**

**4 May 2022**



## 1 Introduction

The COVID-19 pandemic has had a massive effect on economies around the world. Across different countries, millions of workers were furloughed or even lost their jobs as businesses struggled to survive (ny2020). Same situation happened in Australia, due to more restrictions, many businesses closed their doors, while employees were working with less hours or being dismissed by companies. As a result of the continuous “lockdown” periods in 2020, the estimates made by the Australian Bureau of Statistics (ABS2021) concluded that 72% of business generated less revenue and the underemployment rate hit historically high with 13.8% by the end of April, 2020, only one month after the COVID-19 outbreak.

Our research is motivated by the lack of quantitative research on the employment of sub-sectors in Australia, as many studies have focused on the aggregated employment rate. A general problem of aggregated researches is the loss of hierarchical information, which may result in a biased conclusion or “a illusion of employment prosperity”. Thus, a quantitative analysis of the sectoral employment and will overcome this problem, giving us a better scope to evaluate the impact of COVID-19 in Australia.

## 2 Research aim and questions

This research will extend anderson2020 by using data on 87 sub-sectors, instead of 19 they used. I will develop a model for the sub-sectors to evaluate the long run effect and the COVID-19 post-impacts. I will also provide an contrafactual analysis based on the assumption “if there is no pandemic”. subsectoral data will provide us more information, which will assist in getting a better employment dynamics understanding of Australia.

The overall research aim is to provide estimates of subsectoral employment based on historical data. Specifically, my goals are:

1. To build up a time series model of employment in 87 sectors of the Australian economy
2. To use this model to conduct our contrafactual analysis.
3. To use this model to determine which sectors ave the highest impact on employment growth in the long run.

### 3 Review of literature

Our review of literature mainly focuses on two areas:

1. The COVID-19 sectoral impacts and modelling to Economies.
2. Modelling of large numbers of time series.

#### 3.1 Sectoral Impact for COVID-19 to Economies

Existing studies have focused on the evaluation of impacts of COVID-19 on board sectors of large economies such as the US and Europe. **ludvigson2020covid** developed a disaster series to translate the macroeconomic impact of costly and deadly disasters in recent US history and modelled as sectoral shock to predict COVID-19, concluding that the shock would lead to a cumulative loss of 20% in industrial production, 39% in service sector and also reduce the US GDP by 12.75 per cent at the end of 2020. **gregory2020pandemic** conducted simulations under different scenarios via a search theoretic model using US data and found the recovery in US is L-shaped, with employment remaining lower than the pre-covid for a long period. They also extended their studies at disaggregated level of 20 sectors, showing that “arts and entertainment” and “accommodation and food services” sectors would have the biggest shock during the pandemic.

In Australia, **anderson2020** developed a multivariate time series for 19 main sectors in Australia (as a small open economy) using a Bayesian VARX model. Their research concluded that “Manufacturing” and “Construction” have highest positive spillovers for the aggregate economy. Meanwhile, they also applied a “conditional forecasting” method proposed by **waggoner1999** to simulate different scenarios for pandemic in Australia. However, this research does not use the most disaggregated level in Australia (sub-sectors of main sectors), which can be extremely useful in macroeconomic analysis.

#### 3.2 Modelling

##### 3.2.1 Baysian VAR

Literature of multivariate, large-size data tend to rely on Bayesian Vector Autoregression model(BVAR) (**anderson2020**; **litterman1986**; **banbura2010large**). The BVAR model is attractive because it allows us to estimate a large number of parameters, when sample size is not large, in a statistically coherent way.(**litterman1986**; **wozniak2016bayesian**).

In order to utilize the Bayesian VAR estimators and decrease the weight of hte lagged variable with the lag length, **litterman1979** proposed the Minnesota Prior by forcing the means of parameters ‘centered’ as a random walk. The mean on its first own lag is set to unity and the rest are set to zero so that (a) the most recent lag should provide more information than distant lags. (b) own lags should explain more than the lags of other variables.

### 3.2.2 Improvement of BVAR

The literature suggests that a significant improvement can be made in the large BVAR dynamic model by imposing a stronger shrinkage parameter ([banbura2010large](#); [litterman1986](#)). [robertson1999vector](#) and [kadiyala1997](#) proposed a normal inverted Wishart prior which retains the principal of Minnesota prior. Meanwhile, [banbura2010large](#) suggested an easier way to apply the Minnesota prior via adding dummy observations in the BVAR system.

## 4 Data collection and exploratory analysis

### 4.1 Data Introduction

Our data source is the ABS Employment by industry subdivision of main job ([ABS2022](#)), which records employment (measured in '000) by the ANZSIC industry sub-division of their main jobs from 1984 : Q4 to 2021 : Q4. Our data structure is provided via Figure ??.

Although seasonally adjusted data is available in ([ABS2022](#)), I will work with original data to capture any possible changes in seasonal patterns. I an seasonally difference will be applied to the logarithm of the original series in order to make it stationary and eliminate seasonality, which will make it easier to conduct further steps.

### 4.2 Priliminary Exploratory Data Analysis

Main Sub-Sectors with highest and lowest difference in 2020:Q2 compare with Q1

Date	Industry	Top five	Industry	bottom five
2020 Q2	30 Building Construction	30.7	91 Sports and Recreation Activities	-58.8
2020 Q2	81 Tertiary Education	29.6	69 Professional, Scientific and Technical Services (Except Computer System) Design and Related Services)	-59.2
2020 Q2	01 Agriculture	27.0	42 Other Store-Based Retailing	-65.1
2020 Q2	26 Electricity Supply	23.4	89 Heritage Activities	-93.1
2020 Q2	34 Machinery and Equipment Wholesaling	12.9	45 Food and Beverage Services	-255.4

**Figure 1:** The highest and lowest five sub-sectors' employment change in (10000) at 2020:Q2

Figure ?? illustrates the changes of raw data for 19 sectors in a more disaggregated manner. Due to the closedown of businesses and travel bans on 2020:Q2, we can observe that the number dropped substantially (from around 13200 to 12200 on 2020 : Q2). Most industries behaved similarly with significant changes shown in Figure ?? . Comparing with the previous data of these industries, Accommodation&Food, Media&telecom and Administrative industries have experienced a severe loss of employment and have not fully recovered to the pre-covid level. However, industries like Financial and Electricity, Gas show an continuously increase trend as before.

Meanwhile, at 87 sub-sector level (the most disaggregated) in Table ??, we can conclude that the Food and Beverage Service experienced a severe shock after the lockdown, following by Heritage Activities and Other Store-Based Retailing. Figure ?? demonstrates the performance of each industry in the most disaggregated level, we can see sub-sectors we mentioned before have shown huge decreases in employment. It further proves that the employment decreases hugely in those sub-sectors mentioned before.

Nevertheless, there is a draw back of sectoral employment by comparing Figure ?? and Figure ??, that is subsectoral dynamics may not be homogeneous. For some industries in these two figures, the huge changes are driven by some dominant sub-sectors. For example, in 19 sectoral level we may believe that all of the subsectors experienced a large shock by only looking their aggregated performance for Construction sector. (see Figure ??). However, the reality is while Construction sector is decreasing, its sub-sectors 20. Building Construction and 31. Heavy and civil engineering Construction increased (see Figure ??). This means that not all sub-sectors suffered from the COVID-19 even the overall industry got influenced.

## 5 Methodology

### 5.1 Proposed Model

I plan to use a Bayesian VARX model based on a method proposed by **anderson2020**. In the model, each sector is affected by the lags of sectoral growth and a lag of the total employment growth. The lag of aggregate employment growth act as an economy-wide factor which will affect each sector.

Under the assumption that the structure of Australian economy will not change during the COVID-19, we suggest the BVAR model as

$$\mathbf{y}_t = \mathbf{c} + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \mathbf{A}_3\mathbf{y}_{t-3} + \mathbf{A}_4\mathbf{y}_{t-4} + \mathbf{\Gamma}\mathbf{x}_{t-1} + \mathbf{u}_t$$

where  $\mathbf{y}_t$  is an  $87 \times 1$  vector of subsectoral employment growth rate at time  $t$  and  $\mathbf{x}_{t-1}$  is the  $4 \times 1$  vector of 4 lags of the growth rate of aggregate employment (this vector of variables are predetermined at time  $t$ ),  $\mathbf{c}$  is a vector of constants,  $\mathbf{A}_{1,2,3,4}$  are  $87 \times 87$  parameter matrices.  $\mathbf{\Gamma}$  is a  $87 \times 4$  matrix and  $\mathbf{u}_t$  is a vector of reduced form errors with the mean equals to zero and independent variance  $\mathbf{u}_t \sim (\mathbf{0}, \mathbf{\Sigma})$ . (see Appendix)

### 5.2 Prior and shrinkage

I plan to estimate the VARX using Bayesian methods by specifying a Minnesota prior (**anderson2020**; **litterman1986**; **robertson1999vector**). In order to set up the Minnesota prior in our BVAR model, **banbura2010large** suggestes a nature conjugate Normal-Wishart prior that applies shrinkage to the VAR slope coefficients using a Minnesota-type prior.

$$E[a_j^{jk}] = E[\gamma_i^j] = 0$$

$$Var[a_j^{jk}] = \begin{cases} \frac{\lambda^2}{i^2}, & j = k \\ \frac{\lambda^2}{i^2} \frac{\sigma_j^2}{\sigma_k^2}, & otherwise \end{cases}$$

$$Var[\gamma_i^j] = \frac{\lambda^2}{i^2} \frac{\sigma_j^2}{\sigma_e^2}$$

**banbura2010large** stated that the natural conjugate Normal-Inverse-Wishart posterior moments can be calculated either analytically or though adding the dummy observations. I will use dummy observations to estimate the BVAR e.g.(**banbura2010large**; **anderson2020**; **wozniak2016bayesian**). An example completed by **anderson2020** will be given in the Appendix. After the estimation of Bayesian

**Table 1:** *Completed work*

Timeline	Tasks
Week 2	Background reading about Multivariate VAR model
Week 3	Collect Sub-sectoral data from ABS website Prepare materials related to subsectoral employment evaluations
Week 4	Prepare literature review related to BVAR model and setting priors Analysis about the MATLAB code done by Andereson et al. (2020)
Week 5	Read articles about Conditional Forecasting
Week 6	Prepare the conditional forecasting coding and hierarchical forecasting
Week 7	Exploratory data analysis on sectoral and subsectoral data
Week 8	Select suitable plots and test softwares to prepare the proposal
week 9	Write research proposal and prepare the first presentation

**Table 2:** *Research plan*

Timeline	Tasks
June - July	Modelling of subsectoral employment
August	Train the model using pre-covid data with BVAR model
September	apply forecasts for quarters between 2020 to 2022
October	Write my thesis and prepare for the second presentation

VAR and then amounts to conducting least squares, we can draw our VAR coefficients by solving the posterior mode and mean.

## 6 Timeline and expected outcome

Table ?? summarises the work done to date. Table ?? maps out the plan for completing the thesis research.

It is expected that we will have an accurate multivariate BVAR model to forecast the scenario without COVID-19 based on the data of 87 sub-sectors before 2019. Then we have some

## 7 Acknowledgement

I would like to gratitude my supervisor Professor Farshid Vahid and my coordinator Professor Heather Anderson very much for their selfless support and devoted care along the way.

## 8 Appendix

Here is an example of the BVAR estimation done by **anderson2020**.

$$\mathbf{y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{A}_2 + \cdots + \mathbf{A}_4 \mathbf{y}_{t-4} + \mathbf{\Gamma}_1 \mathbf{x}_{t-1} + \mathbf{u}_t \quad (1)$$

$$= \begin{bmatrix} c_1 \\ \vdots \\ c_2 \end{bmatrix} + \begin{bmatrix} a_1^{11} & \cdots & a_1^{1n} & \cdots & a_4^{11} & \cdots & a_4^{1n} & \gamma_1^1 & \cdots & \gamma_4^1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n1} & \cdots & a_1^{nn} & \cdots & a_4^{n1} & \cdots & a_4^{nn} & \gamma_1^n & \cdots & \gamma_4^n \end{bmatrix} \begin{bmatrix} \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-4} \\ x_{t-1} \\ \vdots \\ x_{t-4} \end{bmatrix} \quad (2)$$

$$+ \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{n,t} \end{bmatrix} \quad (3)$$

$$(4)$$

where  $E(\mathbf{u}_t \mathbf{u}_t') = \mathbf{\Sigma}$  and  $E(\mathbf{u}_t \mathbf{u}_{t-1}') = \mathbf{0}$ . Here the  $n$  represent the number of sectors (or in our case will be 87) and  $\mathbf{c}$  represents the vector of constants. There are 4 lags included for the total employment ( $x_{t-1}, \dots, x_{t-4}$ ) as predetermined variable at time  $t$ .

Here is an Minnesota prior example done by **anderson2020**:

The Normal-Wishart prior distribution take the form of

$$\text{vec}(\beta | \mathbf{\Sigma}) \sim N(\text{vec}(\beta_0), \mathbf{\Sigma} \otimes \mathbf{\Omega}_0) \text{ and}$$

$$\mathbf{\Sigma} \sim \text{IW}(S_0, a_0)$$

where we set the prior parameters and  $a_0$  based on the prior setting. The expectation of  $\mathbf{\Sigma}$  being  $\text{diag}(\sigma_1^2, \dots, \sigma_n^2)$ . Then we implement our VAR by defining dummy observations



$$Y_d = \begin{pmatrix} \mathbf{0}_{np+p,n} \\ \mathbf{diag}(\sigma_1, \dots, \sigma_n) \\ \mathbf{0}_{1 \times n} \end{pmatrix}$$

$$X_d = \begin{pmatrix} \mathbf{J}_p \otimes \mathbf{diag}(\frac{\sigma_1}{\lambda} \dots \frac{\sigma_n}{\lambda}, \frac{\sigma_e}{\lambda}) & \mathbf{0}_{(np+p) \times 1} \\ \mathbf{0}_{n,np+p} & \mathbf{0}_{n \times 1} \\ \mathbf{0}_{1,np+p} & \epsilon \end{pmatrix}$$

where

$$\mathbf{J}_p = \mathbf{diag}(1, \dots, p)$$

$$\mathbf{S}_0 = (\mathbf{Y}_d - \mathbf{X}_d \times \mathbf{B}_0)'(\mathbf{Y}_d - \mathbf{X}_d \mathbf{B}_0)$$

$$\mathbf{B}_0 = (\mathbf{X}_d' \mathbf{X}_d)^{-1} \mathbf{X}_d' \mathbf{Y}_d, \quad \mathbf{\Omega}_0 = (\mathbf{X}_d' \mathbf{X}_d)^{-1} \text{ and}$$

$$a_0 = T_d - np - p - 1$$

where  $T_d$  is the number of rows for both  $\mathbf{Y}_d$  and  $\mathbf{X}_d$ .

We can get

$$\mathbf{Y}^* = \mathbf{X}^* \beta + \mu^* \quad \text{where :}$$

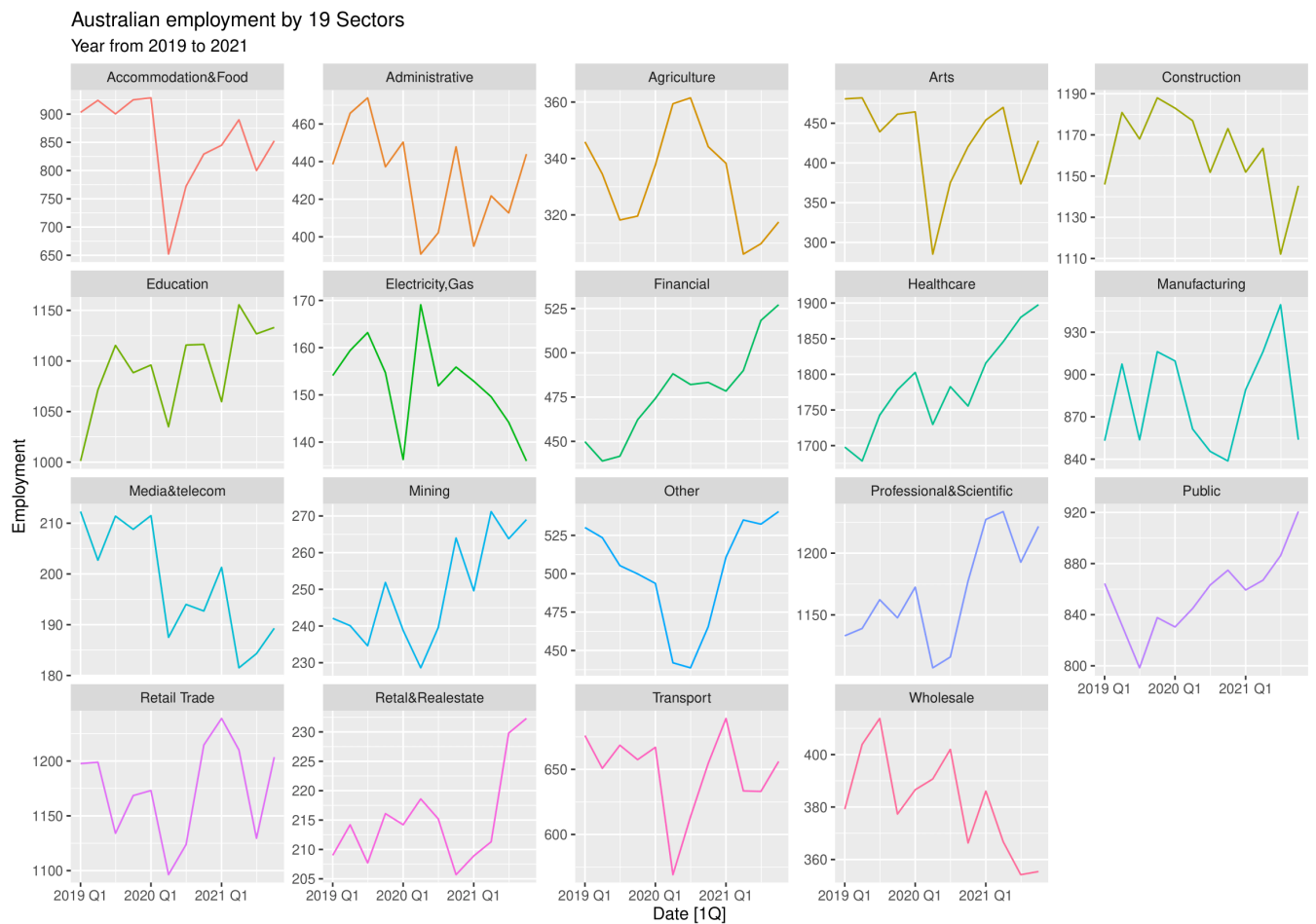
$$\mathbf{Y}^* = [\mathbf{Y}', \mathbf{Y}_d']'; \quad \mathbf{X}^* = [\mathbf{X}', \mathbf{X}_d']'; \quad \mu^* = [\mu', \mu_d']'$$

Then we can estimating the BVAR by conducting an least square regression of  $\mathbf{Y}^*$  on  $\mathbf{X}^*$ . The posterior distribution then have the form of

$$vec(\beta) | \Sigma, Y \sim N(vec(\tilde{\beta}), \Sigma \otimes (\mathbf{X}^{*'} \mathbf{X}^*)^{-1}) \text{ and}$$

$$\Sigma | Y \sim \mathbf{IW}(\tilde{\Sigma}, T_d + T - np + 2)$$

where  $\tilde{\beta} = (\mathbf{X}^{*'} \mathbf{X}^*)^{-1} \mathbf{X}^{*'} \mathbf{Y}^*$  and  $\tilde{\Sigma} = (\mathbf{Y}^* - \mathbf{X}^* \tilde{\beta})'(\mathbf{Y}^* - \mathbf{X}^* \tilde{\beta})$



**Figure 2:** *Employment('000) of 19 sectors in Australia from 2019:Q1 to 2021:Q4*

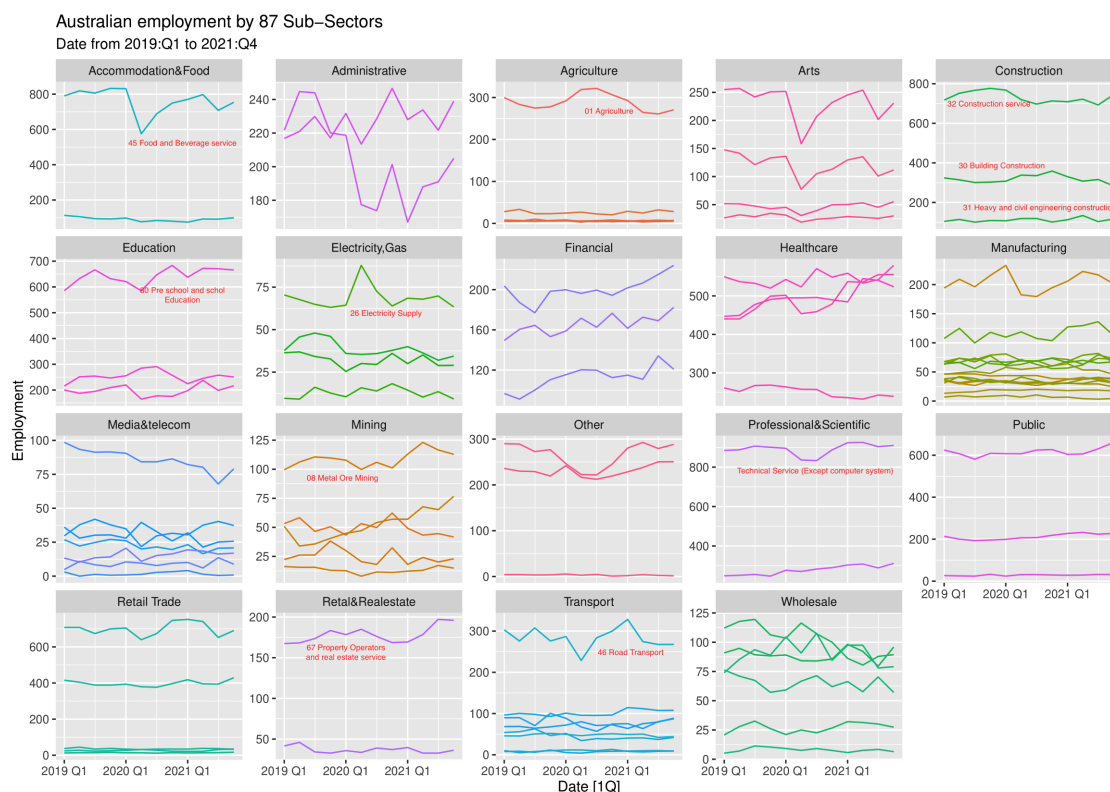


Figure 3: Employment('000) of 87 sub-sectors in Australia from 2019:Q1 to 2021:Q4

## INDUSTRY CLASSIFICATION PYRAMID

ANZSIC [Australian & New Zealand standard industry classification]

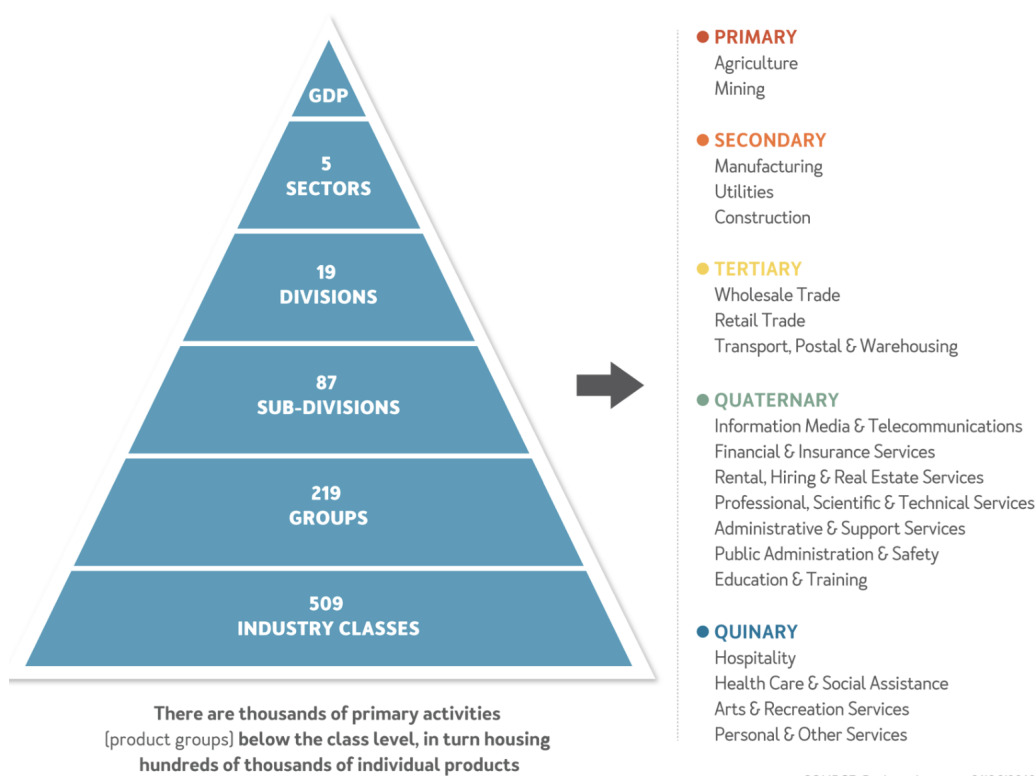


Figure 4: Australian Industry Pamamid plot by (ANZSIC)