Name: Han Hong Tuck from EP0302_01

Title of Data Analysis: How do Global Events influence Singapore's Economic Growth?

Questions to answer to gain deeper insights into dataset:

Question 1: What is the relationship between GDP Growth rate and Contributors to GDP Growth Rate by Industry

Question 2: Why are there fluctations in the GDP Growth rate by industry and how does that influence the overall GDP Growth Rate Annually?

Question 3: What is the trend for the GDP Growth Rate as well as GDP Growth Rate by Industry of the individual datasets?

Url of Annual Contribution to Growth in Gross Domestic Product in Chained 2015 dollars Dataset: https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual (https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual)

Url of Annual Contribution to Growth in Gross Domestic Product in Chained 2015 dollars (By Industry) Dataset: https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual?resource_id=2391f506-3f78-4041-a97d-5ed1c8c48083 (https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual?resource_id=2391f506-3f78-4041-a97d-5ed1c8c48083)

Breaking Down the Nature of GDP Growth Dataset:

The nature of the GDP Growth dataset contains the GDP Growth Rate for Singapore from 1960s to 2010s.

What is Gross Domestic Product (GDP) ?

Gross Domestic Product (GDP) is an important economic indicator of the country's performance over a given period, by measuring the total monetary or market value produced by the country.

Why is Growth in Gross Domestic Product (GDP) important?

Since the growth rate of Gross Domestic Product (GDP) is often used as an indicator of the general health of the economy, an increase in GDP is often interpreted as a sign that the economy is doing well.

What is the metric used to measure the Growth in Gross Domestic Product (GDP)?

The metric used in measuring the Growth in Gross Domestic Product (GDP) for this dataset is using Chained Dollars. Chained dollars is used to adjust the real dollar amounts for inflation over time. When using chained dollars, a base year will be chosen to calculate the actual inflation in the economy, so as to make it fair for comparing the Gross Domestic Product (GDP) Growth across different years. In order to capture the economic information accurately, the base year chosen has to be constantly updated to recent years. The base year also needs to be taken into consideration of the economic environment of the period (relatively stable economic performance), considered somewhere in the middle between high economic performance and low economic performance indicators. Hence, the base year for this dataset is 2015.

Write Python code that uses the pandas package to extract useful statistical or summary information about the data

In [1]:

```
import pandas as pd
import numpy as np

gdp_growth = pd.read_csv("./Datasets/contribution-to-growth-in-gross-domestic-pro

#Shape of GDP Growth Dataset
print(f"Shape of the GDP Growth Dataset: {gdp_growth.shape}\n")

#Index of GDP Growth Dataset
print(f"Index of the GDP Growth Dataset: {gdp_growth.index}\n")

#Column of GDP Growth Dataset
print(f"Columns of the GDP Growth Dataset: \n{gdp_growth.columns}")

#Renaming the column names
gdp_growth.rename(columns=
    {"year":"Financial Year",
     "level_1":"GDP Growth in Chained Dollars (Base Year = 2015)",
     "value":"Percentage Growth (%)"}
    ,inplace=True)

#First 5 rows of GDP Growth Dataset
print("\n\t\tFirst 5 rows of GDP Growth Dataset")
display(gdp_growth.head())

#Last 5 rows of GDP Growth Dataset
print("\n\t\tLast 5 rows of GDP Growth Dataset")
display(gdp_growth.tail())

#Display Summary Information of GDP Growth Dataset
print("\n\t\tSummary infomation of GDP Growth Dataset\n")
display(gdp_growth.info())

#Display top 10 years with fastest GDP Growth
print("\n\t\tTop 10 Years with Fastest GDP Growth in chronological order")
highest_gdp = gdp_growth.sort_values(by="Percentage Growth (%)",ascending=False).

#convert column for GDP Growth to index using pivot
highest_gdp = highest_gdp.pivot(index="GDP Growth in Chained Dollars (Base Year =
display(highest_gdp)

#Display top 10 years with slowest GDP Growth
print("\n\t\tTop 10 Years with Slowest GDP Growth in chronological order")
slowest_gdp = gdp_growth.sort_values(by="Percentage Growth (%)",ascending=False).

#convert column for GDP Growth to index using pivot
slowest_gdp = slowest_gdp.pivot(index="GDP Growth in Chained Dollars (Base Year =
display(slowest_gdp)
```

```
Shape of the GDP Growth Dataset: (58, 3)

Index of the GDP Growth Dataset: RangeIndex(start=0, stop=58, step=1)

Columns of the GDP Growth Dataset:
Index(['year', 'level_1', 'value'], dtype='object')
```

```
                First 5 rows of GDP Growth Dataset
```

	Financial Year	GDP Growth in Chained Dollars (Base Year = 2015)	Percentage Growth (%)
0	1961	GDP In Chained (2015) Dollars	8.1
1	1962	GDP In Chained (2015) Dollars	7.6
2	1963	GDP In Chained (2015) Dollars	10.0
3	1964	GDP In Chained (2015) Dollars	-3.1
4	1965	GDP In Chained (2015) Dollars	7.8

```
                Last 5 rows of GDP Growth Dataset
```

	Financial Year	GDP Growth in Chained Dollars (Base Year = 2015)	Percentage Growth (%)
53	2014	GDP In Chained (2015) Dollars	3.9
54	2015	GDP In Chained (2015) Dollars	2.9
55	2016	GDP In Chained (2015) Dollars	3.0
56	2017	GDP In Chained (2015) Dollars	3.7
57	2018	GDP In Chained (2015) Dollars	3.1

```
                Summary infomation of GDP Growth Dataset

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 58 entries, 0 to 57
Data columns (total 3 columns):
 #   Column                                            Non-Null Count  Dtype  
---  ------                                            --------------  -----  
 0   Financial Year                                    58 non-null     int64  
 1   GDP Growth in Chained Dollars (Base Year = 2015)  58 non-null     object 
 2   Percentage Growth (%)                             58 non-null     float64
dtypes: float64(1), int64(1), object(1)
memory usage: 1.5+ KB
```

```
None
```

```
                Top 10 Years with Fastest GDP Growth in chronological order
```

	Percentage Growth (%)									
Financial Year	**1967**	**1968**	**1969**	**1970**	**1971**	**1972**	**1988**	**1993**	**1994**	**2010**
GDP Growth in Chained Dollars (Base Year = 2015)										
GDP In Chained (2015) Dollars	12.5	13.5	13.8	13.9	12.4	13.3	11.3	11.5	11.1	14.5

```
                Top 10 Years with Slowest GDP Growth in chronological order
```

	Percentage Growth (%)									
Financial Year	1964	1985	1986	1998	2001	2008	2009	2015	2016	2018
GDP Growth in Chained Dollars (Base Year = 2015)										
GDP In Chained (2015) Dollars	-3.1	-0.6	1.3	-2.2	-1.1	1.9	0.1	2.9	3.0	3.1

In [3]:

```
data1 = [1961,1970,1980,1990,2000,2010]
data2 = [1969,1979,1989,1999,2009,2018]
years = 1960

#using loop to display the gdp growth every 10 years
for index, (value1, value2) in enumerate(zip(data1, data2)):
    print(f"\n\t\t\t   GDP Growth in {years}s")
    gdp_growth_every_ten_years = gdp_growth.set_index("Financial Year").loc[value

    #extract statistical info about the gdp every 10 years
    gdp_statistical_info = gdp_growth_every_ten_years.groupby(["GDP Growth in Cha

    #using pivot to convert the column gdp to index
    gdp_growth_every_ten_years = gdp_growth_every_ten_years.pivot(index="GDP Grow

    display(gdp_growth_every_ten_years)

    print(f"\n\t\tStatistical Information for GDP Growth in {years}s")
    display(gdp_statistical_info)
    years += 10
```

```
                            GDP Growth in 1960s
```

	Percentage Growth (%)								
Financial Year	**1961**	**1962**	**1963**	**1964**	**1965**	**1966**	**1967**	**1968**	**1969**
GDP Growth in Chained Dollars (Base Year = 2015)									
GDP In Chained (2015) Dollars	8.1	7.6	10.0	-3.1	7.8	10.2	12.5	13.5	13.8

```
                Statistical Information for GDP Growth in 1960s
```

	Percentage Growth (%)							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
GDP Growth in Chained Dollars (Base Year = 2015)								
GDP In Chained (2015) Dollars	9.0	8.933333	5.105879	-3.1	7.8	10.0	12.5	13.8

```
                            GDP Growth in 1970s
```

	Percentage Growth (%)									
Financial Year	**1970**	**1971**	**1972**	**1973**	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**
GDP Growth in Chained Dollars (Base Year = 2015)										
GDP In Chained (2015) Dollars	13.9	12.4	13.3	10.6	6.1	4.0	7.4	6.9	7.8	9.6

Statistical Information for GDP Growth in 1970s

	Percentage Growth (%)							
	count	mean	std	min	25%	50%	75%	max
GDP Growth in Chained Dollars (Base Year = 2015)								
GDP In Chained (2015) Dollars	10.0	9.2	3.306559	4.0	7.025	8.7	11.95	13.9

GDP Growth in 1980s

	Percentage Growth (%)									
Financial Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
GDP Growth in Chained Dollars (Base Year = 2015)										
GDP In Chained (2015) Dollars	10.1	10.8	7.1	8.6	8.8	-0.6	1.3	10.8	11.3	10.2

Statistical Information for GDP Growth in 1980s

	Percentage Growth (%)							
	count	mean	std	min	25%	50%	75%	max
GDP Growth in Chained Dollars (Base Year = 2015)								
GDP In Chained (2015) Dollars	10.0	7.84	4.16632	-0.6	7.475	9.45	10.65	11.3

GDP Growth in 1990s

	Percentage Growth (%)									
Financial Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
GDP Growth in Chained Dollars (Base Year = 2015)										
GDP In Chained (2015) Dollars	9.8	6.7	6.6	11.5	11.1	7.2	7.5	8.3	-2.2	5.7

Statistical Information for GDP Growth in 1990s

	Percentage Growth (%)							
	count	mean	std	min	25%	50%	75%	max
GDP Growth in Chained Dollars (Base Year = 2015)								
GDP In Chained (2015) Dollars	10.0	7.22	3.843841	-2.2	6.625	7.35	9.425	11.5

GDP Growth in 2000s

	Percentage Growth (%)									
Financial Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
GDP Growth in Chained Dollars (Base Year = 2015)										
GDP In Chained (2015) Dollars	9.0	-1.1	3.9	4.5	9.8	7.4	9.0	9.0	1.9	0.1

```
Statistical Information for GDP Growth in 2000s
```

	Percentage Growth (%)							
	count	mean	std	min	25%	50%	75%	max
GDP Growth in Chained Dollars (Base Year = 2015)								
GDP In Chained (2015) Dollars	10.0	5.35	4.053325	-1.1	2.4	5.95	9.0	9.8

```
GDP Growth in 2010s
```

	Percentage Growth (%)								
Financial Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
GDP Growth in Chained Dollars (Base Year = 2015)									
GDP In Chained (2015) Dollars	14.5	6.3	4.4	4.8	3.9	2.9	3.0	3.7	3.1

```
Statistical Information for GDP Growth in 2010s
```

	Percentage Growth (%)							
	count	mean	std	min	25%	50%	75%	max
GDP Growth in Chained Dollars (Base Year = 2015)								
GDP In Chained (2015) Dollars	9.0	5.177778	3.656767	2.9	3.1	3.9	4.8	14.5

Breaking Down the Nature of GDP Growth by Industry Dataset:

The nature of the GDP Growth by industry contains the contribution of the GDP Growth for Singapore from 1960s to 2010s sorted by their respective industries (Goods Producing Industries, Services Producing Industries, Ownership Of Dwellings, Add: Taxes On Products). These sectors help to contribute the GDP Growth of Singapore's economic development.

- Goods Producing Industries refers to industries from Mandfacturing, Construction, Utilities and other goods industries.
- Services Producing Industries refer to industries from Wholesale and Retail trade, Transportation and Storage, Accomodation and Food services and many more... It typically refers to industries that provide a service instead of end-products like financial consultants or insurance agents.

- Ownership of dwellings refers to revenue generated from Housing services provided by owner-occupiers and individuals who sell their residential properties.
- Taxes on products refers to taxes payable per unit of goods and services when they are produced, delivered, sold, transferred or disposed of by their producers. Examples include goods and services tax (GST), stamp duties and Certificate of Entitlement of motor vehicles.

In [4]:

```
gdp_growth_by_industry = pd.read_csv("./Datasets/contribution-to-growth-in-gdp-in

#Shape of GDP Growth by Industry Dataset
print(f"Shape of the GDP Growth by Industry Dataset: {gdp_growth_by_industry.shap

#Index of GDP Growth by Industry Dataset
print(f"Index of the GDP Growth by Industry Dataset: {gdp_growth_by_industry.inde

#Column of GDP Growth by Industry Dataset
print(f"Columns of the GDP Growth by Industry Dataset: \n{gdp_growth_by_industry.

#Renaming the column names
gdp_growth_by_industry.rename(columns=
    {"year":"Financial Year",
     "level_1":"GDP Growth in Chained Dollars (Base Year = 2015)",
     "level_2":"Type_of_Industries",
     "value":"Percentage Growth (%)"}
    ,inplace=True)

#First 5 rows of GDP Growth by Industry Dataset
print("\n\t\tFirst 5 rows of GDP Growth by Industry Dataset")
display(gdp_growth_by_industry.head())

#Last 5 rows of GDP Growth by Industry Dataset
print("\n\t\tLast 5 rows of GDP Growth by Industry Dataset")
display(gdp_growth_by_industry.tail())

#Display Information of GDP Growth by Industry Dataset
print("\n\t\tSummary infomation of GDP Growth by Industry Dataset\n")
display(gdp_growth_by_industry.info())

#Display Statistical Information of GDP Growth by Industry Dataset
print("\nStatistical infomation of Overall GDP Growth by Industry Dataset\n")
display(gdp_growth_by_industry.groupby(["Type_of_Industries"])[["Percentage Growt

data1 = [1961,1970,1980,1990,2000,2010]
data2 = [1969,1979,1989,1999,2009,2018]
years = 1960

gdp_growth_by_industry.drop(["GDP Growth in Chained Dollars (Base Year = 2015)"],

#using a loop to display the statistical information of gdp growth of all the dat
for index, (value1, value2) in enumerate(zip(data1, data2)):

    print(f"\n\t\t\tGDP Growth in {years}s")
    gdp_growth_every_ten_years = gdp_growth_by_industry.set_index("Financial Year

    #extract statistical info about the gdp every 10 years
    gdp_statistical_info = gdp_growth_every_ten_years.groupby(["Type_of_Industrie

    #using pivot to convert the column gdp to index
    gdp_growth_every_ten_years = gdp_growth_every_ten_years.pivot(index="Type_of_

    display(gdp_growth_every_ten_years)

    print(f"\nStatistical information of GDP Growth by Industry in {years}s")
    display(gdp_statistical_info)
    years += 10
```

```
Shape of the GDP Growth by Industry Dataset: (232, 4)

Index of the GDP Growth by Industry Dataset: RangeIndex(start=0, stop=232, step
=1)

Columns of the GDP Growth by Industry Dataset:
Index(['year', 'level_1', 'level_2', 'value'], dtype='object')

                First 5 rows of GDP Growth by Industry Dataset
```

	Financial Year	GDP Growth in Chained Dollars (Base Year = 2015)	Type_of_Industries	Percentage Growth (%)
0	1961	GDP In Chained (2015) Dollars	Goods Producing Industries	1.8
1	1961	GDP In Chained (2015) Dollars	Services Producing Industries	5.9
2	1961	GDP In Chained (2015) Dollars	Ownership Of Dwellings	0.3
3	1961	GDP In Chained (2015) Dollars	Add: Taxes On Products	0.1
4	1962	GDP In Chained (2015) Dollars	Goods Producing Industries	1.9

```
                Last 5 rows of GDP Growth by Industry Dataset
```

	Financial Year	GDP Growth in Chained Dollars (Base Year = 2015)	Type_of_Industries	Percentage Growth (%)
227	2017	GDP In Chained (2015) Dollars	Add: Taxes On Products	0.3
228	2018	GDP In Chained (2015) Dollars	Goods Producing Industries	1.2
229	2018	GDP In Chained (2015) Dollars	Services Producing Industries	1.9
230	2018	GDP In Chained (2015) Dollars	Ownership Of Dwellings	0.2
231	2018	GDP In Chained (2015) Dollars	Add: Taxes On Products	-0.1

```
                Summary infomation of GDP Growth by Industry Dataset

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 232 entries, 0 to 231
Data columns (total 4 columns):
 #   Column                                           Non-Null Count  Dtype  
---  ------                                           --------------  -----  
 0   Financial Year                                   232 non-null    int64  
 1   GDP Growth in Chained Dollars (Base Year = 2015)  232 non-null    object 
 2   Type_of_Industries                               232 non-null    object 
 3   Percentage Growth (%)                            232 non-null    float6
4
```

```
dtypes: float64(1), int64(1), object(2)
memory usage: 7.4+ KB

None


Statistical infomation of Overall GDP Growth by Industry Dataset
```

	Percentage Growth (%)							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Type_of_Industries								
Add: Taxes On Products	58.0	0.274138	0.354203	-0.9	0.025	0.3	0.475	1.1
Goods Producing Industries	58.0	2.232759	1.852489	-3.6	1.225	2.2	3.400	6.5
Ownership Of Dwellings	58.0	0.208621	0.173998	-0.1	0.100	0.2	0.300	0.8
Services Producing Industries	58.0	4.582759	2.520870	-4.8	3.050	4.9	6.300	8.9

```
                        GDP Growth in 1960s
```

	Percentage Growth (%)								
Financial Year	**1961**	**1962**	**1963**	**1964**	**1965**	**1966**	**1967**	**1968**	**1969**
Type_of_Industries									
Add: Taxes On Products	0.1	0.3	0.0	0.2	1.1	0.7	0.0	0.0	0.8
Goods Producing Industries	1.8	1.9	2.7	1.4	2.3	2.9	3.4	4.6	4.6
Ownership Of Dwellings	0.3	0.3	0.2	0.1	0.2	0.3	0.2	0.4	0.3
Services Producing Industries	5.9	5.1	7.1	-4.8	4.2	6.3	8.9	8.5	8.1

```
Statistical information of GDP Growth by Industry in 1960s
```

	Percentage Growth (%)							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Type_of_Industries								
Add: Taxes On Products	9.0	0.355556	0.409607	0.0	0.0	0.2	0.7	1.1
Goods Producing Industries	9.0	2.844444	1.165237	1.4	1.9	2.7	3.4	4.6
Ownership Of Dwellings	9.0	0.255556	0.088192	0.1	0.2	0.3	0.3	0.4
Services Producing Industries	9.0	5.477778	4.162565	-4.8	5.1	6.3	8.1	8.9

```
                        GDP Growth in 1970s
```

	Percentage Growth (%)									
Financial Year	**1970**	**1971**	**1972**	**1973**	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**

Type_of_Industries	Percentage Growth (%)									
Financial Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Type_of_Industries										
Add: Taxes On Products	0.9	0.2	0.7	0.4	-0.3	0.0	0.1	0.3	0.2	0.5
Goods Producing Industries	5.3	4.7	5.2	2.6	1.2	1.0	3.7	2.1	2.2	3.8
Ownership Of Dwellings	0.4	0.2	0.4	0.1	0.2	0.1	0.1	0.2	0.1	0.2
Services Producing Industries	7.3	7.4	7.1	7.4	5.0	2.7	3.7	4.3	5.3	5.1

```
Statistical information of GDP Growth by Industry in 1970s
```

	Percentage Growth (%)							
	count	mean	std	min	25%	50%	75%	max
Type_of_Industries								
Add: Taxes On Products	10.0	0.30	0.346410	-0.3	0.125	0.25	0.475	0.9
Goods Producing Industries	10.0	3.18	1.587311	1.0	2.125	3.15	4.475	5.3
Ownership Of Dwellings	10.0	0.20	0.115470	0.1	0.100	0.20	0.200	0.4
Services Producing Industries	10.0	5.53	1.699052	2.7	4.475	5.20	7.250	7.4

```
                          GDP Growth in 1980s
```

	Percentage Growth (%)									
Financial Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Type_of_Industries										
Add: Taxes On Products	0.4	0.2	0.3	0.3	0.2	0.0	0.0	0.5	0.4	0.8
Goods Producing Industries	3.3	3.7	1.8	3.6	3.4	-3.6	-0.6	3.2	4.2	2.7
Ownership Of Dwellings	0.1	0.6	-0.1	0.8	0.8	0.4	0.4	0.3	0.2	0.3
Services Producing Industries	6.3	6.2	5.2	3.9	4.4	2.5	1.5	6.7	6.4	6.4

```
Statistical information of GDP Growth by Industry in 1980s
```

	Percentage Growth (%)							
	count	mean	std	min	25%	50%	75%	max
Type_of_Industries								
Add: Taxes On Products	10.0	0.31	0.237814	0.0	0.200	0.30	0.400	0.8
Goods Producing Industries	10.0	2.17	2.443608	-3.6	2.025	3.25	3.550	4.2
Ownership Of Dwellings	10.0	0.38	0.289828	-0.1	0.225	0.35	0.550	0.8
Services Producing Industries	10.0	4.95	1.826502	1.5	4.025	5.70	6.375	6.7

```
                    GDP Growth in 1990s
```

	Percentage Growth (%)									
Financial Year	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**
Type_of_Industries										
Add: Taxes On Products	0.4	0.0	0.4	0.7	0.4	0.4	0.3	0.5	-0.5	0.4
Goods Producing Industries	2.9	2.2	1.6	3.0	4.6	3.0	2.0	2.0	0.1	2.2
Ownership Of Dwellings	0.0	0.1	0.0	0.1	0.2	0.1	0.1	0.2	0.4	0.4
Services Producing Industries	6.5	4.4	4.6	7.7	6.0	3.7	5.0	5.6	-2.1	2.7

```
Statistical information of GDP Growth by Industry in 1990s
```

	Percentage Growth (%)							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Type_of_Industries								
Add: Taxes On Products	10.0	0.30	0.329983	-0.5	0.325	0.4	0.400	0.7
Goods Producing Industries	10.0	2.36	1.160651	0.1	2.000	2.2	2.975	4.6
Ownership Of Dwellings	10.0	0.16	0.142984	0.0	0.100	0.1	0.200	0.4
Services Producing Industries	10.0	4.41	2.694212	-2.1	3.875	4.8	5.900	7.7

```
                    GDP Growth in 2000s
```

	Percentage Growth (%)									
Financial Year	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Type_of_Industries										
Add: Taxes On Products	0.6	-0.3	0.1	0.3	0.6	0.2	0.4	0.8	-0.9	0.1
Goods Producing Industries	3.4	-2.9	1.5	0.5	3.3	2.6	3.6	2.2	-0.2	0.1
Ownership Of Dwellings	0.3	0.1	0.2	0.0	0.1	0.1	0.1	0.0	0.0	0.1
Services Producing Industries	4.8	2.0	2.2	3.7	5.8	4.4	5.0	6.0	2.9	-0.1

```
Statistical information of GDP Growth by Industry in 2000s
```

	Percentage Growth (%)							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Type_of_Industries								
Add: Taxes On Products	10.0	0.19	0.495424	-0.9	0.100	0.25	0.550	0.8
Goods Producing Industries	10.0	1.41	2.057210	-2.9	0.200	1.85	3.125	3.6
Ownership Of Dwellings	10.0	0.10	0.094281	0.0	0.025	0.10	0.100	0.3

Percentage Growth (%)

```
GDP Growth in 2010s
```

	Percentage Growth (%)								
Financial Year	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Type_of_Industries									
Add: Taxes On Products	0.3	-0.1	0.2	-0.2	0.0	0.6	0.7	0.3	-0.1
Goods Producing Industries	6.5	1.8	0.5	0.5	0.9	-0.6	0.6	1.3	1.2
Ownership Of Dwellings	0.0	0.1	0.2	0.1	0.2	0.2	0.2	0.2	0.2
Services Producing Industries	7.8	4.5	3.5	4.4	2.8	2.7	1.4	1.9	1.9

```
Statistical information of GDP Growth by Industry in 2010s
```

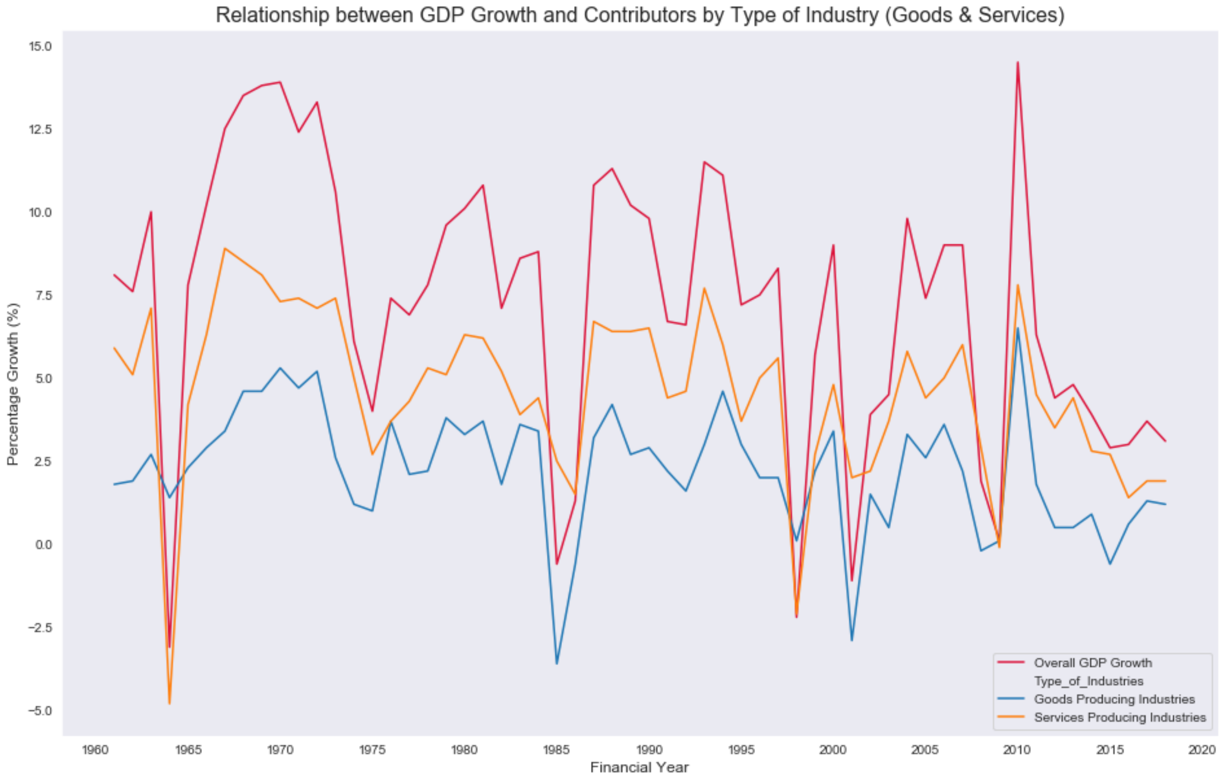
	Percentage Growth (%)							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Type_of_Industries								
Add: Taxes On Products	9.0	0.188889	0.317980	-0.2	-0.1	0.2	0.3	0.7
Goods Producing Industries	9.0	1.411111	2.022650	-0.6	0.5	0.9	1.3	6.5
Ownership Of Dwellings	9.0	0.155556	0.072648	0.0	0.1	0.2	0.2	0.2
Services Producing Industries	9.0	3.433333	1.965960	1.4	1.9	2.8	4.4	7.8

Write Python code that uses Python Visualisation Package (Matplotlib/Seaborn..) to produce useful data visualizations that explain the data.

Data Visualisation 1: Comparative Line Graph (Seaborn & Matplotlib)

In [5]:

```
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib.ticker as ticker

sns.set_style("dark")

gdp_growth = pd.read_csv("./Datasets/contribution-to-growth-in-gross-domestic-pro

#drop column which specify every row as gdp growth rate
gdp_growth.drop(["level_1"],axis=1,inplace=True)

#declaring figure and axes to plot
fig,ax = plt.subplots(figsize=(16,10))

gdp_growth_by_industry = pd.read_csv("./Datasets/contribution-to-growth-in-gdp-in

#renaming the column name
gdp_growth_by_industry.rename(columns=
    {"year":"Financial Year",
     'level_2':"Type_of_Industries",
     "value":"Percentage Growth (%)"}
    ,inplace=True)

gdp_growth_by_industry = gdp_growth_by_industry[gdp_growth_by_industry.Type_of_In
gdp_growth_by_industry.drop(["level_1"],axis=1,inplace=True)

#plotting overall gdp_growth using matplotlib
ax.plot(gdp_growth,label="Overall GDP Growth",color="crimson")

#plotting gdp_growth for goods and services producing industry using seaborn
sns.lineplot(x='Financial Year', y='Percentage Growth (%)', hue='Type_of_Industri

#displaying legend for the individual lines
ax.legend()

#set the frequency of the xticks and yticks on the x-axis and y-axis
ax.xaxis.set_major_locator(ticker.MultipleLocator(5))

#set the title of the graph
ax.set_title("Relationship between GDP Growth and Contributors by Type of Industr

#setting label and fontsize for x_axis and y_axis
plt.xlabel('Financial Year', fontsize=12),plt.ylabel("Percentage Growth (%)",font

#show the figure
plt.show()
```



For the chosen datasets, explain the nature of that dataset (i.e. what is in that dataset) or any pecularities about it you wish to highlight and explain the process you went through to analyse that dataset, . Where possible, you should specifically mention how you used the Pandas, Matplotlib, Seaborn functions to achieve a certain outcome e.g. to transform the data or to produce a certain visualization:

Process of using Pandas, Matplotlib or Seaborn functions to transform the data:

Using Pandas to transform GDP growth dataset

For the GDP Growth dataset, I first rename the column names of the dataset using the rename method. Then, I used the shape method to see the number of rows and columns. After that, I call the method .index and .columns to get the index range and column names of the dataset. Then, I used the head and tail method to see the first and last 5 rows of dataset. I also want to retrieve summary information of the dataset so I apply the info method of the dataset to determine the data type of the columns. I was interested in the years whether singapore is experiencing the fastest gdp growth so I sort the percentage growth by descending order and calling the head method. Similarly, I was interested in the years where singapore experienced the slowest gdp growth so I sort the dataset in descending order and calling the tail method.

Besides that, I am also interested to learn about the average gdp growth every ten years from 1960s to 2010s as well the individual gdp growth during that period. Hence, I used an enumerate and zip function to loop every 10 years to calculate and display the average gdp growth for that period as well using the pivot to convert the column gdp to index to display the gdp growth for that period (e.g in 1960s).

Using Pandas to transform GDP growth by industry dataset

For the GDP Growth by industry dataset, most of the methods that were applied at the start were similar to the GDP Growth Dataset but I make a few changes to the last few methods. I am interested to understand the statistcial information for the gdp contribution in individual industry so I applied the describe method on the groupby - sorted by the type of industries for all the data. However, I am still curious to learn about statistical information for the period every 10 years for the individual industries so I used a loop to enumerate and display the information using the groupby to sort by type of industries.

Using Matplotlib and/or Seaborn functions to transform the data:

Firstly, I import matplotlib and seaborn and read the data from the excel for gdp growth and gdp growth by industry. Then I drop the level_1 column for the gdp growth dataset which refers the individual rows as the gdp growth as I already knew that the dataset is about the gdp growth rate in chained dollars with the base year as 2015. Then I declare fig and ax object for plotting and increase the figure size by passing parameter (16:8). After I have done that, I rename the columns of the gdp growth by industry so that it will easier to understand the column names.

After that, I filter the gdp growth by industry only for the goods producing industry and services producing industry as they are a strong indicator for fluctauations in the gdp growth rate. This is because after extracting the statistical information for all four industries ("Goods Producing Industries" ,"Services Producing Industries","Ownership Of Dwellings","Add: Taxes On Products") using the pandas package, I found out that contribution to gdp for ownership of dwellings and taxes on products always hover at the same growth rate between 2% and -2%. Hence, It is difficult to establish the relationship between the GDP Growth rate for that year as well as the contributors to gdp growth rate for these two respective industries.

Therefore, I specified to only filter for the goods and services producing industries as they offer a much better picture of the relationship between them. After doing that, I reset the index and drop the level 1 for the gdp growth dataset as it is no longer needed. After that, I plot the gdp growth rate and the contributors to the gdp growth rate for goods and services industry using matplotlib for the gd growth rate and seaborn lineplot for the contributors. Then I called the method ax.legend and set the xtick frequency to period of 5.

For each dataset, highlight the insights you have gained from analysing the data and any conclusions or recommendations you want to make as a result of the analysis:

To answer the first question, I am able to tell that there is a very strong correlation between the contributors to the GDP Growth Rate as well as the anually GDP Growth Rate. This is evident as from the graph, I am able to tell that when the overall gdp growth rate decreases, the GDP Contribution for the goods producing and services producing industries decreases as well.

Moving on on the second question, I am able to tell that from the graph, there are a period of years that experience very strong fluctuations such as the year 1964,1986,1998 and 2001. In order to understand why the GDP Figures fluctuate a lot from these period of time, there is need to know how global events influence singapore's economic performance.

Singapore is one of the Southeast Asia countries and it relies heavily on strong trade and financial linkages with the region. In 1998, Singapore was experiencing an Asian Financial Crisis which was caused by currency devaluations of the thailand's currency. This has resulted in a large portion of asian currencies fell and a reduce in demand for bilaterial trade by countries affected by it.

With that decrease in demand, the goods producing industries in Singapore suffered a sharp decline for trade demands and this has also indirectly affected the services producing industries in Singapore.As such, Singapore is very vulnerable to economic events that happen around us and that contributes to why Singapore's GDP figures constantly fluctuates.

Lastly, I am able to infer that as Singapore approach the year 1990s to 2010s, it is experiencing more fluctuations in the GDP Growth rate compared to the period from 1960s to 1990s. This is evident as between 1960s and 1990s, Singapore only experienced two major fluctuations in 1964 and 1985. However, from the year 1990s to 2010s, I can tell that Singapore three major sharp declines in 1998,2001 and 2008.

One of the reasons why Singapore is experiencing this is because during the early years of Singapore's development, Singapore is still at the early stages of economic progress and the economic development is still not intertwined deeply with foreign countries.

As Singapore continues to develop over the years, Singapore's has established itself as a financial and trade hub, and a growing presence in global economy. As such, Singapore being an open economic environment, it is more subceptible to uncertainty due to global economic events. Therefore, it is experiencing more fluctuations in the GDP Growth than before.

Name: Han Hong Tuck from EP0302_01

Title of Data Analysis: How do Global Events influence Singapore's Economic Growth?

Questions to answer to gain deeper insights into dataset:

Question 1: How does global demands influence Nature of Retrenchments in Singapore's Economy, given the fluctuations in Contributions to GDP Growth?

Question 2: How much impact does global events on the Nature of Retrenchments in Singapore?

Question 3: Is there a strong correlation between the GDP Growth Rate and Nature of Retrenchments in Singapore

Url of Annual Contribution to Growth in Gross Domestic Product in Chained 2015 dollars Dataset: https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual (https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual)

Url of Annual Retrenched Employees (By Industry) Dataset: https://data.gov.sg/dataset/retrenched-employees-by-industry-and-occupational-group-annual?resource_id=21c9e4c1-1524-47db-b8c0-59ad89df2a6f (https://data.gov.sg/dataset/retrenched-employees-by-industry-and-occupational-group-annual?resource_id=21c9e4c1-1524-47db-b8c0-59ad89df2a6f)

Url of Annual Short Work-Week & Temporary Layoffs (By Industry) Dataset: https://data.gov.sg/dataset/short-work-week-temporary-layoffs-annual?resource_id=65c5e391-d2a9-43b3-8004-9287e0936521 (https://data.gov.sg/dataset/short-work-week-temporary-layoffs-annual?resource_id=65c5e391-d2a9-43b3-8004-9287e0936521)

Breaking Down the Nature of Retrenchments in Jobs (By Industry) Dataset:

The nature of the dataset for retrenched employees contains the retrenchment of permanent jobs, retrenchments of Term Contract (Temporary) Jobs as well as total retrenchments from the year 1998 to 2019 due to redundancy, sorted by the various industries (manufacturing, construction, services, others).

Defining Terminologies:

1. Retrenchment refers to the termination of permanent employees due to redundancy and early termination of term contract employees due to redundancy. Permanent Jobs typically do not have a predetermined end date to employment.

2. Retrenchment of Term Contract Employees refers to early termination of term contract employees due to redundancy. Term Contract (Temporary) Jobs refers Jobs that are fixed-term contracts which can be terminated by employers upon expiry of a specific term or period (such as a date), unless it is renewed.
3. Total Retrenchments refers to Retrenchments inclusive of Permanent Jobs as well as Term Contract (Temporary) Jobs.

Redundancy refers refers to a situation whereby an employer reduces their workforce in the event that a certain jobs are no longer needed. Such situations may arise due to exterior factors due to the global economic enviroment which has resulted in businesses closing down, the employer needing to cut expenses, the advent of artificial technology (AI) or other technologies that have made that job unnecessary.

Defining Various Industries:

- The Manufacturing Industry comprises mainly of Electronics, Chemicals, Biomedical Sciences, Logistics and Transport engineering.
- The Construction Industry comprises mainly of Construction of Singapore Residential and Non-Residential Buildings.
- The Services Industry comprises mainly of Wholesale & Retail Trade, Recreation, Community & Personal Services and Business services.
- Others refers to Industries that do not belong to any of specific sectors listed in Manufacturing, Construction and Services Industries.

Why is the Number of Job Retrenchments by Industry important?

During a recession, there will be a significant decline in economic activity spreading across the various industries in Singapore, due to a lack of demand for goods and services caused by global environment economic factors. As a result, business and firms in Singapore may be forced to reallocate resources, scale back production and limit losses and as a result, retrench their employees.

Therefore, the number of retrenchments in Singapore is a very important economic indicator of how badly impacted the various type of industries. These data helps the government in considering how they are able to better design and implement policies to help the badly-affected employees who were retrenched due to the economic situation in Singapore.

Breaking Down the Nature of Jobs affected by Short Work Week and Temporary Layoff (By Industry) Dataset:

The nature of the dataset for short work week and temporary layoff contains the number of jobs that implemented short work week, the number of jobs that have been temporary laid off as well as the total number of jobs from the year 1998 to 2019, sorted by the various industries (manufacturing, construction, services, others).

Defining Terminologies:

1. Short Work-Week are employees whose normal number of working days per week has been temporarily reduced due to lack of work, at any time during the reference period.

2. Temporary Layoffs are employees whose services are suspended temporarily due to lack of work, at any time during the reference period. They may or may not have been paid during this period.

Defining Various Industries:

- The Manufacturing Industry comprises mainly of Electronics, Chemicals, Biomedical Sciences, Logistics and Transport engineering.
- The Construction Industry comprises mainly of Construction of Singapore Residential and Non-Residential Buidlings.
- The Services Industry comprises mainly of Wholesale & Retail Trade, Recreation, Community & Personal Services and Business services.
- Others refers to Industries that do not belong to any of specific sectors listed in Manufacturing, Construction and Services Industries.

Why is the Nature of Jobs affected by Short Work Week and Temporary Layoff by Industry important?

During a recession, there will be a significant decline in economic activity spreading across the various industries in Singapore, due to a lack of demand for goods and services caused by global environment economic factors. As a result, business and firms in Singapore may be forced to reallocate resources, scale back production and limit losses and as a result, retrench their employees.

Some businesses and firms may consider implementing alternative measures as they would like to keep their businesses viable while still supporting their employees during period of economic downturns, reducing the number of jobs losses. In liew of such economic environment, businesses and firms will be able to learn how to better handle excessive manpower so that while achieving their objective of restructuring their businesses, the employees' interest will still not be compromised.

These not only reduces the number of job losses due to retrenchments, it also enhances the business and firm's individual competiveness in the industry. This information also provides a strong gauge of the type of industries that adopt alternatives to retrenchments instead of immediate termination of employees.

Write Python code that uses the pandas package to extract useful statistical or summary information about the data

In [2]:

```
import pandas as pd
import numpy as np

retrenchments_by_industry = pd.read_csv("./Datasets/retrenchment-by-industry-leve
short_work_week_and_temporary_layoffs_by_industry = pd.read_csv("./Datasets/short

#Merging retrenchments dataframe with short work week and temporary layoff datafr
df_all = pd.merge(retrenchments_by_industry, short_work_week_and_temporary_layoff

#Shape of Retrenchments, Short Work Week and Temporary Layoffs by Industry Datase
print(f"Shape of Retrenchments, Short Work Week and Temporary Layoffs by Industry

#Index of Retrenchments, Short Work Week and Temporary Layoffs by Industry Datase
print(f"Index of Retrenchments, Short Work Week and Temporary Layoffs by Industry

#Column of Retrenchments, Short Work Week and Temporary Layoffs by Industry Datas
print(f"Columns of Retrenchments, Short Work Week and Temporary Layoffs by Indust

#Renaming the column names
df_all.rename(
    columns=
    {"year":"Year",
     "industry1":"Type_of_Industry",
     "retrench":"Total Retrenchments",
     "retrench_permanent": "Retrenchment of Permanent Jobs",
     "retrench_term_contract":"Retrenchment of Term Contract (Temporary) Jobs",
     "total": "Short Work Week and Temporary Layoffs (Total)",
     "short_work_week": "Short Work Week",
     "temporary_layoff":"Temporary Layoff"}
    ,inplace=True)

#First 5 rows of Retrenchments, Short Work Week and Temporary Layoffs by Industry
print("\n\t\tFirst 5 rows of Retrenchments, Short Work Week and Temporary Layoffs
display(df_all.head())

#Last 5 rows of Retrenchments, Short Work Week and Temporary Layoffs by Industry
print("\n\t\tLast 5 rows of Retrenchments, Short Work Week and Temporary Layoffs
display(df_all.tail())

#Display Summary information of Retrenchments, Short Work Week and Temporary Layo
print("\n\t\tSummary Information of Retrenchments, Short Work Week and Temporary
display(df_all.info())

#Display Statistical Information of Retrenchments, Short Work Week and Temporary
print("\nStatistical infomation of Retrenchments, Short Work Week and Temporary L
list_of_variables = ["Total Retrenchments", "Retrenchment of Permanent Jobs", "Re
for variable in list_of_variables:
    display(df_all.groupby(["Type_of_Industry"])[[variable]].describe(percentiles

df_all = df_all.set_index("Year").drop(2019).reset_index()
```

```
Shape of Retrenchments, Short Work Week and Temporary Layoffs by Industry Dat
aset:
(88, 8)
```

```
Index of Retrenchments, Short Work Week and Temporary Layoffs by Industry Dat
aset:
RangeIndex(start=0, stop=88, step=1)

Columns of Retrenchments, Short Work Week and Temporary Layoffs by Industry D
ataset:
Index(['year', 'industry1', 'retrench', 'retrench_permanent',
       'retrench_term_contract', 'total', 'short_work_week',
       'temporary_layoff'],
      dtype='object')


                First 5 rows of Retrenchments, Short Work Week and Temporary
Layoffs Dataset
```

	Year	Type_of_Industry	Total Retrenchments	Retrenchment of Permanent Jobs	Retrenchment of Term Contract (Temporary) Jobs	Short Work Week and Temporary Layoffs (Total)	Short Work Week	Tempo La
0	1998	manufacturing	20700.0	18930.0	1770.0	3940.0	3230.0	7
1	1998	construction	2450.0	1210.0	1240.0	90.0	70.0	
2	1998	services	9630.0	8940.0	690.0	420.0	400.0	
3	1998	others	30.0	0.0	30.0	0.0	0.0	
4	1999	manufacturing	8370.0	7990.0	380.0	660.0	570.0	

```
                Last 5 rows of Retrenchments, Short Work Week and Temporary Lay
offs Dataset
```

	Year	Type_of_Industry	Total Retrenchments	Retrenchment of Permanent Jobs	Retrenchment of Term Contract (Temporary) Jobs	Short Work Week and Temporary Layoffs (Total)	Short Work Week	Temp L
83	2018	others	0.0	0.0	0.0	0.0	0.0	
84	2019	manufacturing	2790.0	2580.0	220.0	490.0	480.0	
85	2019	construction	860.0	500.0	370.0	110.0	80.0	
86	2019	services	7000.0	6700.0	300.0	290.0	250.0	
87	2019	others	30.0	30.0	0.0	0.0	0.0	

```
                Summary Information of Retrenchments, Short Work Week and Tem
porary Layoffs Dataset
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 88 entries, 0 to 87
Data columns (total 8 columns):
 #   Column                                         Non-Null Count  Dtype
```

```
---  ------                                          --------------  -----
 0   Year                                            88 non-null     int64
 1   Type_of_Industry                                88 non-null     object
 2   Total Retrenchments                             88 non-null     float64
 3   Retrenchment of Permanent Jobs                  88 non-null     float64
 4   Retrenchment of Term Contract (Temporary) Jobs  88 non-null     float64
 5   Short Work Week and Temporary Layoffs (Total)   88 non-null     float64
 6   Short Work Week                                 88 non-null     float64
 7   Temporary Layoff                                88 non-null     float64
dtypes: float64(6), int64(1), object(1)
memory usage: 5.6+ KB

None


Statistical infomation of Retrenchments, Short Work Week and Temporary Layoffs
Dataset Sorted By Type
```

	Total Retrenchments								
	count	**mean**	**std**	**min**	**20%**	**40%**	**50%**	**60%**	**80%**
Type_of_Industry									
construction	22.0	1044.090909	607.106009	70.0	540.0	842.0	940.0	1030.0	1622.0
manufacturing	22.0	7373.181818	4484.219300	2570.0	4132.0	5084.0	5890.0	7320.0	9500.0
others	22.0	70.909091	89.597464	0.0	10.0	20.0	30.0	50.0	98.0
services	22.0	6748.636364	2428.196069	2990.0	4150.0	6042.0	6640.0	7156.0	8874.0

	Retrenchment of Permanent Jobs								
	count	**mean**	**std**	**min**	**20%**	**40%**	**50%**	**60%**	**80%**
Type_of_Industry									
construction	22.0	597.272727	309.979751	20.0	374.0	482.0	535.0	672.0	840.0
manufacturing	22.0	6768.636364	4092.023944	2280.0	3724.0	4806.0	5435.0	7062.0	8656.0
others	22.0	61.818182	88.351913	0.0	2.0	10.0	20.0	42.0	80.0
services	22.0	6265.000000	2364.796880	2390.0	3930.0	5350.0	6315.0	6644.0	8336.0

	Retrenchment of Term Contract (Temporary) Jobs									
	count	**mean**	**std**	**min**	**20%**	**40%**	**50%**	**60%**	**80%**	**max**
Type_of_Industry										
construction	22.0	448.636364	351.639635	30.0	134.0	308.0	350.0	406.0	780.0	1240.0
manufacturing	22.0	605.000000	563.659387	90.0	226.0	310.0	405.0	496.0	700.0	2030.0
others	22.0	7.272727	21.861429	0.0	0.0	0.0	0.0	0.0	8.0	100.0
services	22.0	484.545455	230.872615	110.0	254.0	444.0	505.0	550.0	682.0	1010.0

	Short Work Week and Temporary Layoffs (Total)									
	count	mean	std	min	20%	40%	50%	60%	80%	max
Type_of_Industry										
construction	22.0	109.545455	112.227544	0.0	30.0	54.0	85.0	100.0	168.0	41
manufacturing	22.0	2025.454545	3122.137130	130.0	434.0	616.0	655.0	930.0	2076.0	1137
others	22.0	1.818182	5.010811	0.0	0.0	0.0	0.0	0.0	0.0	2
services	22.0	500.000000	731.729849	30.0	76.0	140.0	185.0	246.0	748.0	278

	Short Work Week									
	count	mean	std	min	20%	40%	50%	60%	80%	max
Type_of_Industry										
construction	22.0	70.454545	84.935677	0.0	12.0	34.0	40.0	60.0	104.0	31
manufacturing	22.0	1838.181818	2869.492466	120.0	376.0	546.0	610.0	902.0	1848.0	1073
others	22.0	1.818182	5.010811	0.0	0.0	0.0	0.0	0.0	0.0	2
services	22.0	445.000000	644.225264	20.0	66.0	118.0	145.0	212.0	696.0	225

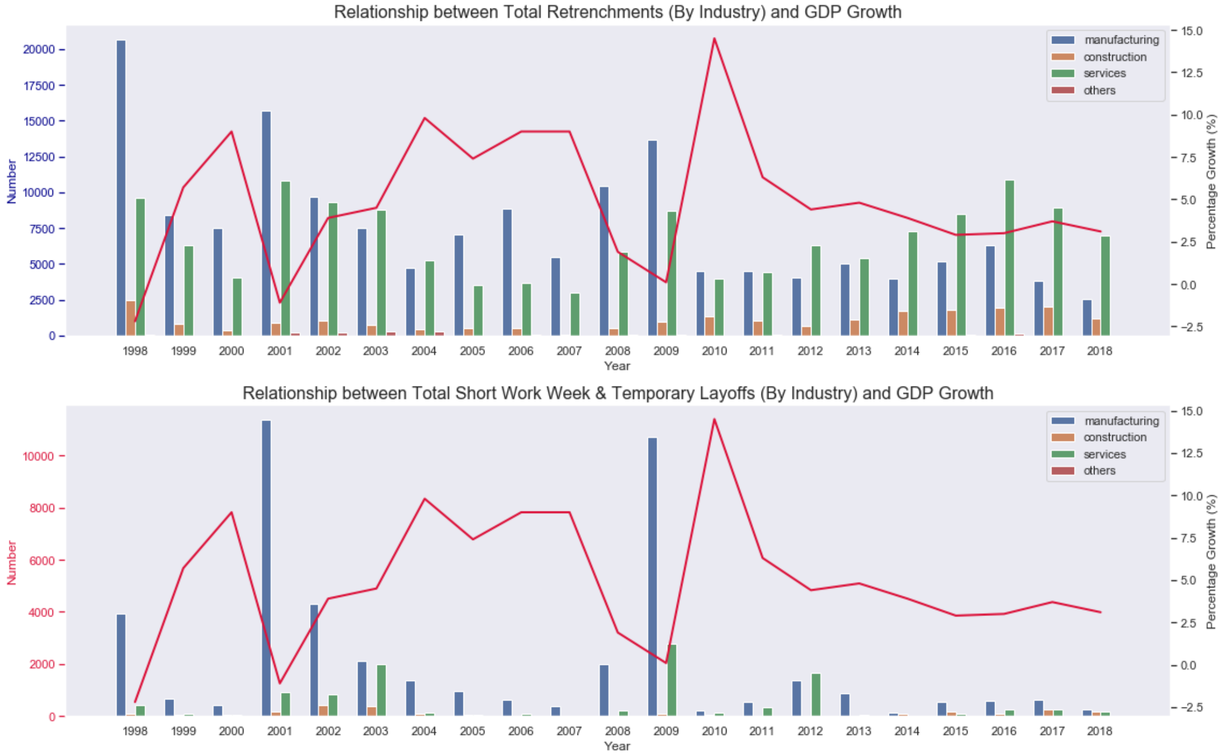
	Temporary Layoff									
	count	mean	std	min	20%	40%	50%	60%	80%	max
Type_of_Industry										
construction	22.0	40.909091	38.285779	0.0	10.0	20.0	30.0	36.0	70.0	140.0
manufacturing	22.0	187.727273	283.866699	10.0	20.0	34.0	40.0	108.0	252.0	1090.0
others	22.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0
services	22.0	54.545455	111.343624	0.0	10.0	20.0	30.0	30.0	40.0	530.0

Write Python code that uses Python Visualisation Package (Matplotlib/Seaborn..) to produce useful data visualizations that explain the data.

Data Visualisation 2: Grouped Bar Graphs with Line Graphs (Seaborn & Matplotlib)

In [8]:

```
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib.ticker as ticker

#set style of Seaborn Graph background to dark
sns.set(style="dark")

#read from csv file
gdp_growth = pd.read_csv("./Datasets/contribution-to-growth-in-gross-domestic-pro

#Renaming the column names
gdp_growth.rename(columns=
    {"year":"Financial Year",
     "value":"Percentage Growth (%)"}
    ,inplace=True)

#dropping level 1 columns which identifies each row as the GDP Growth Rate
gdp_growth.drop(["level_1"],axis=1,inplace=True)

#declare fig and axes object to plot
fig,ax = plt.subplots(2,1,figsize=(16,10))

#First Graph for Retrenchments

#convert column for total retrenchments to type and Column values to row values
df_melt_retrenchment = df_all.melt(value_vars=df_all[["Total Retrenchments"]],var

#declare ax0 to plot on the first subplot
ax0 = plt.subplot(211)

#using seaborn barplot function to plot retrenchments by type of industry
sns.barplot(x='Year', y='Number', hue='Type_of_Industry', data=df_melt_retrenchme

#display legend
ax0.legend()

#creating a twin axes to plot gdp growth
ax1 = ax0.twinx()

#using seaborn lineplot to plot gdp growth
sns.lineplot(ax0.get_xticks(), gdp_growth["Percentage Growth (%)"].loc[1998:2018]

#set title for first graph plotting
ax0.set_title("Relationship between Total Retrenchments (By Industry) and GDP Gro

#Second Graph for total short work week and temporary layoff

#convert column for total short work week and temporary layoff to type and Column
df_melt_short_work_week_and_temporary_layoff = df_all.melt(value_vars=df_all[["Sh

#declare ax2 to plot on the first subplot
ax2 = plt.subplot(212)

#using seaborn barplot function to plot total short work week and temporary layof
sns.barplot(x='Year', y='Number', hue='Type_of_Industry', data=df_melt_short_work
#display legend
ax2.legend()

#creating a twin axes to plot gdp growth
ax3 = ax2.twinx()

#using seaborn lineplot to plot gdp growth
sns.lineplot(ax2.get_xticks(), gdp_growth["Percentage Growth (%)"].loc[1998:2018]

#set title for first graph plotting
ax2.set_title("Relationship between Total Short Work Week & Temporary Layoffs (By

#Setting color to differentiate both graphs

#setting tick params to darkblue and crimson to differentiate between both graphs
ax0.tick_params(axis='y', colors='darkblue'), ax2.tick_params(axis='y', colors='c

#setting x and y labels to darkblue and crimson to differentiate between both gra
ax0.yaxis.label.set_color('darkblue'), ax2.yaxis.label.set_color('crimson')

#to display both graph together nicely
fig.tight_layout()

#display graph
plt.show()
```

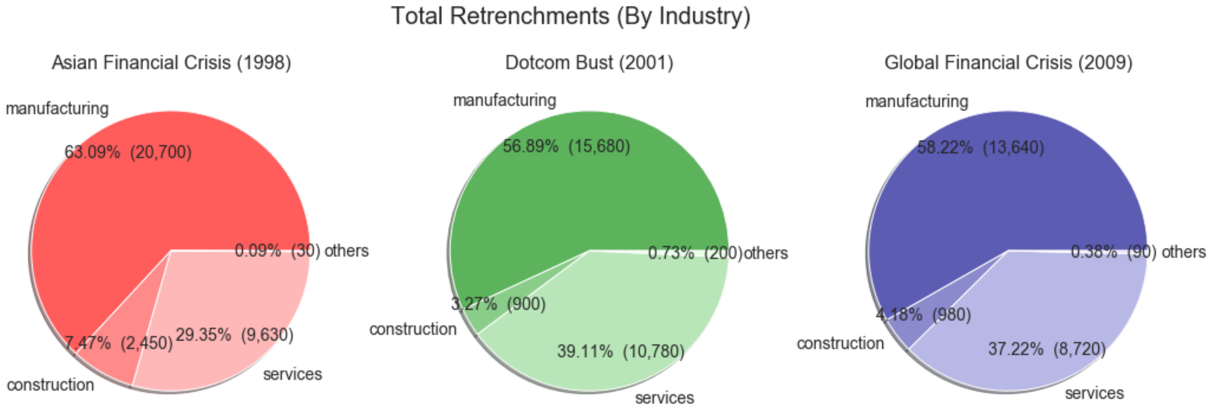


Data Visualisation 3: Pie Charts (Matplotlib)

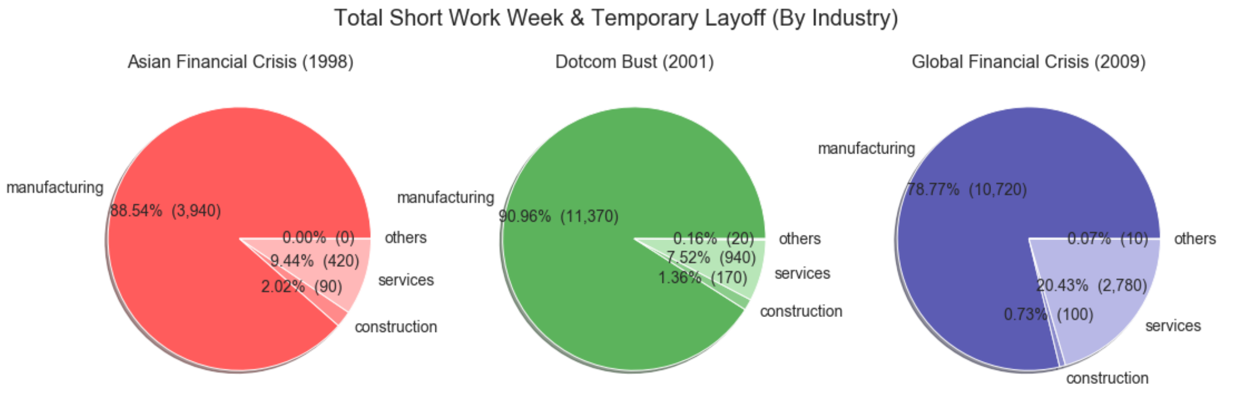
Given the high number of Retrenchments and Short Work Week & Temporary Layoff in 1998, 2001, 2009, these years will be picked from the dataframe for analysing the industries most affected as well as the underlying reasons behind it.

In [3]:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

#set index to year, extract years where there are high retrenchments, drop the ty
df_retrenchment = df_melt_retrenchment.set_index("Year").loc[["1998","2001","2009

#declare fig and axes object for three different pie charts
fig, axes = plt.subplots(1,3,figsize=(18,6))

#declare list for color of individual years, title and count
color = [["#ff5c5c","#ff8a8a","#ffb8b8","#ffe5e5"],["#5cb35c","#8acc8a","#b8e6b8"
title = ["Asian Financial Crisis", "Dotcom Bust", "Global Financial Crisis"]
count = 0

#using zip and for loop to iterate and plot all three pie charts in one figure
for ax,(groupname,subdf) in zip(axes,df_retrenchment.groupby('Year')):
    values = subdf.Number.tolist()
    ax.pie(subdf.Number, labels=subdf.Type_of_Industry, autopct=lambda p : '{:.2f
    ax.set_title(f"{title[count]} ({groupname})",fontsize=16)
    count += 1

#set title above all three pie charts
fig.suptitle("Total Retrenchments (By Industry)",fontsize=20)

#display graph
plt.show()
```



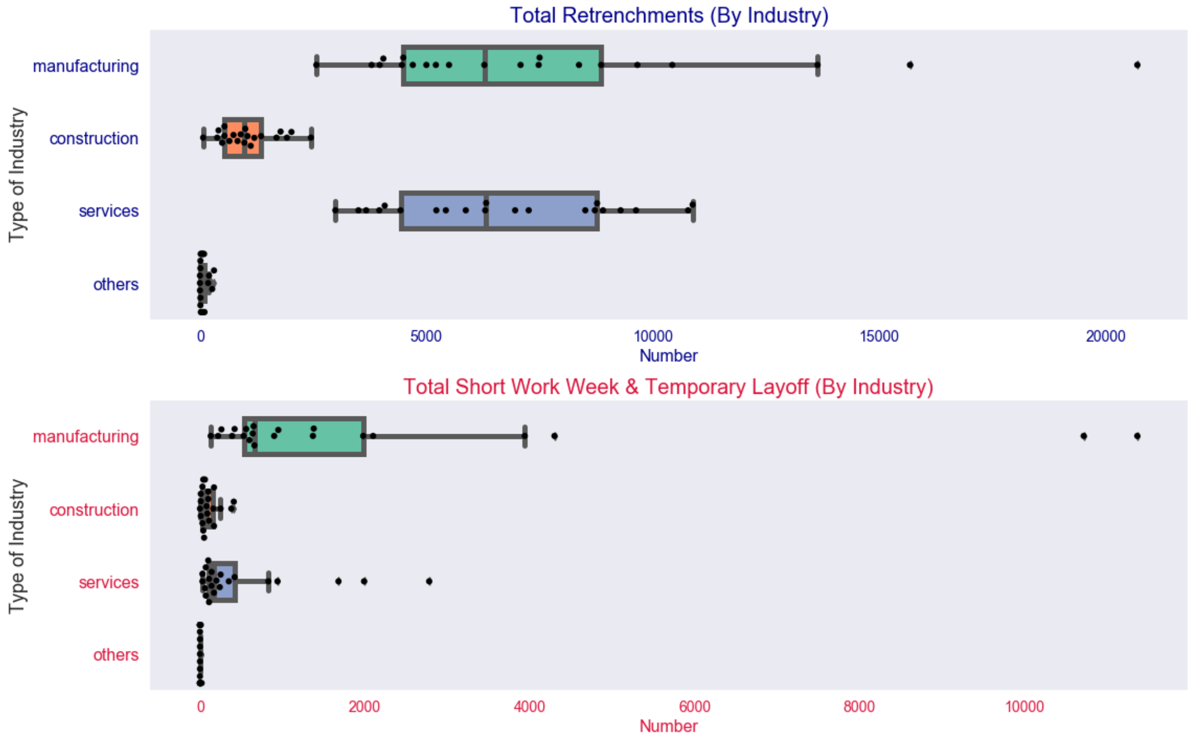
In [4]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

#set index to year, extract years where there are high short work week and tempoe
df_short_work_week_and_temporary_layoff = df_melt_short_work_week_and_temporary_l

#declare fig and axes object for three different pie charts
fig, axes = plt.subplots(1,3,figsize=(18,6))

#declare list for color of individual years, title and count
color = [["#ff5c5c","#ff8a8a","#ffb8b8","#ffe5e5"],["#5cb35c","#8acc8a","#b8e6b8'
title = ["Asian Financial Crisis", "Dotcom Bust", "Global Financial Crisis"]
count = 0

#using zip and for loop to iterate and plot all three pie charts in one figure
for ax,(groupname,subdf) in zip(axes,df_short_work_week_and_temporary_layoff.grou
    values = subdf.Number.tolist()
    ax.pie(subdf.Number, labels=subdf.Type_of_Industry, autopct=lambda p : '{:.2f
    ax.set_title(f"{title[count]} ({groupname})",fontsize=16)
    count += 1

#set title above all three pie charts
fig.suptitle("Total Short Work Week & Temporary Layoff (By Industry)",fontsize=20

#display graph
plt.show()
```



Data Visualisation 4: Boxplots & Swarmplots (Seaborn and Matplotlib)

In [12]:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

#declare fig and axes object to plot
fig, ax = plt.subplots(2,1,figsize=(16,10))

#declare ax0 to plot on the first subplot
ax0 = plt.subplot(211)

#plotting boxplot graph
sns.boxplot(x="Number", y="Type_of_Industry", data=df_melt_retrenchment, width=0.

#plotting swarmplot graph on top of boxplot graph
sns.swarmplot(x="Number", y="Type_of_Industry", data=df_melt_retrenchment, color=

#set title for first subplot
ax0.set_title("Total Retrenchments (By Industry)", fontsize=20, color='darkblue')


#declare ax1 to plot on the second subplot
ax1 = plt.subplot(212)

#plotting boxplot graph
sns.boxplot(x="Number",y="Type_of_Industry", data=df_melt_short_work_week_and_tem

#plotting swarmplot graph on top of boxplot graph
sns.swarmplot(x="Number", y="Type_of_Industry", data=df_melt_short_work_week_and_

#set title for second subplot
ax1.set_title("Total Short Work Week & Temporary Layoff (By Industry)", fontsize=


#set fontsize, color for x labels, y labels, x ticks and y ticks
ax0.set_ylabel('Type of Industry', fontsize=18), ax1.set_ylabel('Type of Industry

ax0.set_xlabel('Number', fontsize=16, color='darkblue'), ax1.set_xlabel('Number',

ax0.tick_params(axis="x", labelsize=14, colors='darkblue'), ax1.tick_params(axis=

ax0.tick_params(axis="y", labelsize=16, colors='darkblue'), ax1.tick_params(axis=

#to display both graph nicely
fig.tight_layout()

#display graph
plt.show()
```



For the chosen datasets, explain the nature of that dataset (i.e. what is in that dataset) or any pecularities about it you wish to highlight and explain the process you went through to analyse that dataset, . Where possible, you should specifically mention how you used the Pandas, Matplotlib, Seaborn functions to achieve a certain outcome e.g. to transform the data or to produce a certain visualization:

Pecularities to highlight:

Duriny years where the number of retrenchments are high, the number of total short work week and temporary layoff is also quite high as well. However, there were significantly lower short work week and layoffs in 1998 (Global Financial Crisis) compared to year 2001 (Dotcom Bust) and 2009 (Global Financial Crisis).

Process of using Pandas, Matplotlib or Seaborn functions to transform the data:

Using Pandas to transform the data

Firstly, I read the excel file from retrenchments and short work week and layoffs and store it in two variables. Then, I merge both dataframes into one and store it as df_all. After that, I display the shape, index and column names of the df_all dataset so that I am able to rename the column names. Moving on, I used the head and tail method to display the first and last five rows. To gather summary information of the data, I used the info method. To display Statistical Information of Retrenchments, Short Work Week and Temporary Layoffs in the df_all dataset, I used a for loop to iterate the list of variable and display the statistics using the groupby function and describe method with varying percentiles.

Using Matplotlib or Seaborn functions to transform the data:

Data Visualisation 2: Grouped Bar Graph with Line Graph

Firstly, I set style of Seaborn Graph background to dark and read from the gdp growth excel file. After that, I rename the column names for gdp growth and dropped level 1 columns which identifies each row as the GDP Growth Rate.

In order to make a comparison between the retrenchments, short work week and temporary layoff and the GDP Growth, I created another variable to read from the gdp growth excel file. Then, I rename the column names and drop level 1 columns which indentifies each row as the GDP Growth Rate. Then, I declare the fig and axes object such that both graphs will display on the same figure, using subplots.

Afterwards, I convert the column for total retrenchments to type and Column values to row values, declare ax0 to plot on the first subplot, and use seaborn barplot function to plot retrenchments by type of industry and display the legend. Then, I called twin axes to declare share another y axis to plot the gdp growth using seaborn lineplot function and setting the title. Similarily, the process of plotting the total short work week and layoff is the same. In order to differentiate both graphs, I specify the colors for both graphs to crimson and darkblue for yticks and labels for y-axis. Then, I call fig tight layout so that both graph will not interfere one another and plt show to display the graph.

Data Visualisation 3: Pie Charts

As pie charts is only able to show the distribution for the retrenchments/short work week and layoff based on individual years, only the number of years where there are high number of the values are selected from the df_retrenchments and df_short_work_week and temporary layoff dataset. The years are 1998,2001 and 2009. First, I set the index to year, extract years where there are high retrenchments rates, drop the type and reset index from the df_retrenchments dataset. Then, I declare fig and axes object for three different pie charts and list for color of individual years, title and count. Then, I use zip and for loop to iterate and plot all three pie charts in one figure set the title above the three pie charts. Similarly, the process of plotting total short work week and temporary layoff is the same but just with different variables.

Data Visualisation 4: Boxplots & Swarmplots

Firstly, I declare fig and axes object to plot. Then, I declare ax0 on the first subplot and plot the graph using seaborn boxplot and swarmplot using df_melt_retrenchment dataframe. Then, I set the title for the first subplot. Similarly, the process to plot total short work week and temporary layoff is the same but on different subplots and using different variables. In order to differentiate between both graph, the color and fontsize of retrenchments and short work week & layoff is set to darkblue and crimson. Then, I call fig tight layout so that both graph will not interfere one another and plt show to display the graph.

For each dataset, highlight the insights you have gained from analysing the data and any conclusions or recommendations you want to make as a result of the analysis:

Data Visualisation 2: Grouped Bar Graph with Line Graph

From the graph, we are able to tell that there is a very strong correlation between total retrenchments, short week and layoff when comparing to GDP Growth. During years where there are low retrenchments rates and short week and layoff, Singapore is experiencing the fastest GDP Growth. This is evidently seen from the year of 2010 after the global financial crisis, which Singapore experiences the fastest GDP Growth and the lowest retrenchments and short work week and layoff. During times where there are high number of retrenchments and short work week and layoff, the GDP Growth rate is inevitably the lowest in year 1998, 2001 and 2009.

Besides that, we are also able to infer that manufacturing industry is always one of the hardest hit during an economic downturn. This can be seen from the graph where the total number of retrenchments during economic crisis for the manufacturing sector is always higher compared to the services and construction sector, with the number of retrenchments being around 20 thousand during 1998 Asian Financial Crisis, 15 thousand during 2001 dotcom bust and 12 thousand during 2009 Global Financial Crisis. The number of short work week and layoffs is also one of the highest during that period, at around 4000, 11000 and 10000 respectively.

Data Visualisation 3: Pie Charts

Given the high number of Retrenchments and Short Work Week & Temporary Layoff in 1998, 2001, 2009, these years will be used to further analysed the industries that are most affected.

After analysing the graph, I are able to tell that the retrenchments for the manufacturing sector is always the hardest hit, followed by the service sector and construction sector.

To understand why manufacturing sectors is always has the highest number of retrenchments, we will use 2001 dotcom bust as an example. The manufacturing sector has always been one of the greatest contributors to GDP Growth in Singapore, accounting to about 20-25% between 1990s and 2000s. The electronics sector in particular has always been the largest contributor to the manufacturing output of Singapore's economy. During the year leading up to 2001 dotcom bubble, there is excessive speculation of technology companies over the potential technology boom and popularizing use of the internet in the US. This sector contributed close to two-thirds of Singapore's non-oil domestic exports (products that were manufactured in Singapore and sold to foreigh countries) back then. As a result, when the technology bubble finally burst in 2001, there were massive retrenchments for people that work in the electronics sector, who were responsible for producing and manufacturing semiconductors, mobile products and computer chips.

Besides that, I am also able to infer that due to the nature of demand of the various industries in Singapore, companies tend to retrench employees instead of employing alternatives to retrenchments such as total short work week and layoff. This can be seen when during the asian financial crisis, the number of retrenchments is at 20000 while alternatives to retrenchments (short work week and layoff) is only around 4000. However, there seems to be increasing awareness of responsible employer practices as the total number of short work week and layoff has increased during economic downturn and the number of retrenchments has decreased. This could be due to the fact that the Singapore Government (Ministry Of Manpower) encourages and provide support for employers to retain and provide guidelines for managing excessive manpower.

Data Visualisation 4: Boxplots & Swarmplots

Besides retrenchments and alternatives to retrenchments due to economic crises, Employers also tend to retrench or layoff their workers due to other reasons as well. In order to better understand the nature of retrenchments and alternative to retrenchments, the boxplot is used to show the distribution of total retrenchments by Industry as well as total short work week and layoff by industry.

Given the nature that employers tend to only impose short work week and layoffs during economic downturn, it is not surprising to see that the various industries tend to have close to zero number of total short work week and layoff. Therefore, plotting boxplot to find the distribution of total short work and layoffs is not a wise decision as there are many outliers. However, I still plotted the boxplot so as to show the comparison between the distribution of total retrenchments and distribution of total short work week and layoffs.

Moving on to retrenchments, economic downturn is not the only reason why employers retrench their workers. There are several reasons why Companies retrench their employees. One of the possible reasons is due to the ageism. The middle aged and older workers tend to be vulnerable to retrenchments as there is a pervasive assumption in Singapore that they are less productive regardless of their actual abilities.

Companies tend to retrench older workers because they are deemed to be less flexible, outdated, not technologically orientated and much more expensive hires compared to younger workers. Hence, there is a wider distribution of the data for retrenchments sorted by the various industry.

An interesting thing to note is that the mean (average) retrenchment rates for the manufacturing sector is the same as the services sector. This could be attributed to the fact that advancements in technologies has resulted in automation and robots replacing job in the service industry.

Therefore, there is always a degree of retrenchments in Singapore.

Name: Han Hong Tuck from EP0302_01

Title of Data Analysis: How do Global Events influence Singapore's Economic Growth?

Questions to answer to gain deeper insights into dataset:

Question 1: Does the Scatter plot futher reaffirms the relationship between GDP Growth, retrenchments and short work week and layoffs?

Question 2: Coming to the end of the Data Analysis, what are some of the key takeaways after analysing all these datasets?

Url of Annual Contribution to Growth in Gross Domestic Product in Chained 2015 dollars Dataset: https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual (https://data.gov.sg/dataset/contribution-to-growth-in-gdp-in-chained-2015-dollars-by-industry-ssic-2015-annual)

Url of Annual Retrenched Employees (Topline): https://data.gov.sg/dataset/retrenched-employees-by-industry-and-occupational-group-annual (https://data.gov.sg/dataset/retrenched-employees-by-industry-and-occupational-group-annual)

Url of Annual Short Work-Week & Temporary Layoffs (Topline): https://data.gov.sg/dataset/short-work-week-temporary-layoffs-annual (https://data.gov.sg/dataset/short-work-week-temporary-layoffs-annual)

Breaking Down the Nature of Retrenchments in Jobs (Topline) Dataset:

The nature of the dataset for retrenched employees contains the retrenchment of permanent jobs, retrenchments of Term Contract (Temporary) Jobs as well as total retrenchments from the year 1998 to 2019 due to redundancy, sorted by year.

Defining Terminologies:

1. Retrenchment refers to the termination of permanent employees due to redundancy and early termination of term contract employees due to redundancy. Permanent Jobs typically do not have a predetermined end date to employment.
2. Retrenchment of Term Contract Employees refers to early termination of term contract employees due to redundancy. Term Contract (Temporary) Jobs refers Jobs that are fixed-term contracts which can be terminated by employers upon expiry of a specific term or period (such as a date), unless it is renewed.
3. Total Retrenchments refers to Retrenchments inclusive of Permanent Jobs as well as Term Contract (Temporary) Jobs.

Redundancy refers refers to a situation whereby an employer reduces their workforce in the event that a certain jobs are no longer needed. Such situations may arise due to exterior factors due to the global economic enviroment which has resulted in businesses closing down, the employer needing to cut expenses, the advent of artificial technology (AI) or other technologies that have made that job unnecessary.

Purpose of Plotting Nature of Retrenchments in Jobs (Topline) Dataset

Since the Nature of Retrenchment in Jobs (By Industry) has been already plotted, the plotting of the Nature of Retrenchments in Jobs (Topline) is to further display the correlation between GDP Growth and Total Retrenchments over the years. By doing so, it will further prove that the hypothesis, analysis and insights gained from analysing the dataset is well-substantiated and justified.

Breaking Down the Nature of Jobs affected by Short Work Week and Temporary Layoff (Topline) Dataset:

The nature of the dataset for short work week and temporary layoff contains the number of jobs that implemented short work week, the number of jobs that have been temporary laid off as well as the total number of jobs from the year 1998 to 2019, sorted by year.

Defining Terminologies:

1. Short Work-Week are employees whose normal number of working days per week has been temporarily reduced due to lack of work, at any time during the reference period.
2. Temporary Layoffs are employees whose services are suspended temporarily due to lack of work, at any time during the reference period. They may or may not have been paid during this period.

Purpose of Nature of Jobs affected by Short Work Week and Temporary Layoff (Topline) Dataset:

Since the Nature of Jobs affected by Short Work Week and Temporary Layoff (By Industry) has been already plotted, the plotting of the Nature of Jobs affected by Short Work Week and Temporary Layoff (Topline) is to further display the correlation between GDP Growth and Total Short Work Week and Layoffs over the years. By doing so, it will further prove that the hypothesis, analysis and insights gained from analysing the dataset is well-substantiated and justified.

Write Python code that uses the pandas package to extract useful statistical or summary information about the data

In [2]:

```
import pandas as pd
import numpy as np

#read from file
retrenchments = pd.read_csv("./Datasets/retrenchment-topline.csv",na_values=["-"]
short_work_week_and_temporary_layoffs = pd.read_csv("./Datasets/short-work-week-a

#Merging retrenchments dataframe with short work week and temporary layoff datafr
df_all = pd.merge(retrenchments, short_work_week_and_temporary_layoffs, on=['year

#Shape of Retrenchments, Short Work Week and Temporary Layoffs Dataset
print(f"Shape of Retrenchments, Short Work Week and Temporary Layoffs by Industry

#Index of Retrenchments, Short Work Week and Temporary Layoffs Dataset
print(f"Index of Retrenchments, Short Work Week and Temporary Layoffs by Industry

#Column of Retrenchments, Short Work Week and Temporary Layoffs Dataset
print(f"Columns of Retrenchments, Short Work Week and Temporary Layoffs by Indust

#Renaming the column names
df_all.rename(
    columns=
    {"year":"Year",
     "industry1":"Type_of_Industry",
     "retrench":"Total Retrenchments",
     "retrench_permanent": "Retrenchment of Permanent Jobs",
     "retrench_term_contract":"Retrenchment of Term Contract (Temporary) Jobs",
     "total": "Short Work Week and Temporary Layoffs (Total)",
     "short_work_week": "Short Work Week",
     "temporary_layoff":"Temporary Layoff"}
    ,inplace=True)

#First 5 rows of Retrenchments, Short Work Week and Temporary Layoffs Dataset
print("\n\t\tFirst 5 rows of Retrenchments, Short Work Week and Temporary Layoffs
display(df_all.head())

#Last 5 rows of Retrenchments, Short Work Week and Temporary Layoffs Dataset
print("\n\t\tLast 5 rows of Retrenchments, Short Work Week and Temporary Layoffs
display(df_all.tail())

#Display Summary information of Retrenchments, Short Work Week and Temporary Layo
print("\n\t\tSummary Information of Retrenchments, Short Work Week and Temporary
display(df_all.info())

#Display Statistical Information of Retrenchments, Short Work Week and Temporary
print("\n\t\tStatistical infomation of Retrenchments, Short Work Week and Tempora
display(df_all.set_index("Year").describe(percentiles = [.20,.40,.60,.80]).T)
```

```
Shape of Retrenchments, Short Work Week and Temporary Layoffs by Industry Datas
et:
(22, 7)

Index of Retrenchments, Short Work Week and Temporary Layoffs by Industry Datas
et:
RangeIndex(start=0, stop=22, step=1)
```

```
Columns of Retrenchments, Short Work Week and Temporary Layoffs by Industry Dat
aset:
Index(['year', 'retrench', 'retrench_permanent', 'retrench_term_contract',
       'total', 'short_work_week', 'temporary_layoff'],
      dtype='object')


                First 5 rows of Retrenchments, Short Work Week and Temporary La
yoffs Dataset
```

	Year	Total Retrenchments	Retrenchment of Permanent Jobs	Retrenchment of Term Contract (Temporary) Jobs	Short Work Week and Temporary Layoffs (Total)	Short Work Week	Temporary Layoff
0	1998	32800	29090	3720	4450	3690	760
1	1999	15530	14620	910	800	680	120
2	2000	11950	11620	330	540	370	170
3	2001	27570	25840	1730	12500	11730	760
4	2002	20130	19090	1040	5540	4980	570

```
                Last 5 rows of Retrenchments, Short Work Week and Temporary Lay
offs Dataset
```

	Year	Total Retrenchments	Retrenchment of Permanent Jobs	Retrenchment of Term Contract (Temporary) Jobs	Short Work Week and Temporary Layoffs (Total)	Short Work Week	Temporary Layoff
17	2015	15580	13440	2140	810	660	150
18	2016	19170	16810	2360	950	820	130
19	2017	14720	12880	1840	1160	980	180
20	2018	10730	9610	1120	590	480	110
21	2019	10690	9810	880	890	820	80

```
                Summary Information of Retrenchments, Short Work Week and Tem
porary Layoffs Dataset

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 22 entries, 0 to 21
Data columns (total 7 columns):
 #   Column                                          Non-Null Count  Dtype
---  ------                                          --------------  -----
 0   Year                                            22 non-null     int64
 1   Total Retrenchments                             22 non-null     int64
 2   Retrenchment of Permanent Jobs                22 non-null     int64
 3   Retrenchment of Term Contract (Temporary) Jobs  22 non-null     int64
 4   Short Work Week and Temporary Layoffs (Total)   22 non-null     int64
 5   Short Work Week                                 22 non-null     int64
 6   Temporary Layoff                                22 non-null     int64
dtypes: int64(7)
memory usage: 1.3 KB

None


                Statistical infomation of Retrenchments, Short Work Week and Te
mporary Layoffs Dataset
```

	count	mean	std	min	20%	40%	50%	60%	8
Total Retrenchments	22.0	15236.363636	6197.475327	8590.0	10698.0	11716.0	13010.0	15206.0	1878
Retrenchment of Permanent Jobs	22.0	13693.636364	5645.694985	7680.0	9698.0	10688.0	12110.0	13216.0	1672
Retrenchment of Term Contract (Temporary) Jobs	22.0	1544.090909	929.088950	330.0	864.0	1034.0	1230.0	1694.0	212
Short Work Week and Temporary Layoffs (Total)	22.0	2637.272727	3681.842424	260.0	624.0	910.0	965.0	1120.0	417
Short Work Week	22.0	2355.454545	3341.938028	150.0	504.0	820.0	915.0	992.0	355
Temporary Layoff	22.0	282.727273	378.671636	20.0	70.0	114.0	140.0	166.0	43

In [3]:

```
#convert column for total retrenchments to type and Column values to row values
df_melt_retrenchment = df_all.melt(value_vars=df_all[["Total Retrenchments"]],var

#Display top 5 years with Highest Retrenchment Rates
print("\nTop 5 years with Highest Total Number of Retrenchments")
highest_retrenchments = df_melt_retrenchment.sort_values(by="Number",ascending=Fa
#convert column for type to index using pivot
display(highest_retrenchments.pivot(index="Type",columns = "Year"))

#Display top 5 years with Lowest Retrenchment Rates
print("\nTop 5 years with Lowest Total Number of Retrenchments")
lowest_retrenchments = df_melt_retrenchment.sort_values(by="Number",ascending=Fal
#convert column for type to index using pivot
display(lowest_retrenchments.pivot(index="Type",columns = "Year"))

#convert column for total short work week and temporary layoff to type and Column
df_melt_short_work_week_and_temporary_layoff = df_all.melt(value_vars=df_all[["Sh

#Display top 5 years with Highest Total Number of Short Work Week and Temporary L
print("\nTop 5 years with Highest Total Number of Short Work Week and Temporary L
highest_retrenchments = df_melt_short_work_week_and_temporary_layoff.sort_values(
#convert column for type to index using pivot
display(highest_retrenchments.pivot(index="Type",columns = "Year"))

#Display top 5 years with Lowest Total Number of Short Work Week and Temporary La
print("\nTop 5 years with Lowest Total Number of Short Work Week and Temporary La
lowest_retrenchments = df_melt_short_work_week_and_temporary_layoff.sort_values(b
#convert column for type to index using pivot
display(lowest_retrenchments.pivot(index="Type",columns = "Year"))
```

```
Top 5 years with Highest Total Number of Retrenchments
```

	Number				
Year	**1998**	**2001**	**2002**	**2009**	**2016**
Type					
Total Retrenchments	32800	27570	20130	23430	19170

```
Top 5 years with Lowest Total Number of Retrenchments
```

	Number				
Year	**2004**	**2007**	**2010**	**2011**	**2019**
Type					
Total Retrenchments	10640	8590	9800	9990	10690

```
Top 5 years with Highest Total Number of Short Work Week and Temporary Layoffs
```

	Number				
Year	**1998**	**2001**	**2002**	**2003**	**2009**
Type					
Short Work Week and Temporary Layoffs (Total)	4450	12500	5540	4470	13610

```
Top 5 years with Lowest Total Number of Short Work Week and Temporary Layoffs
```

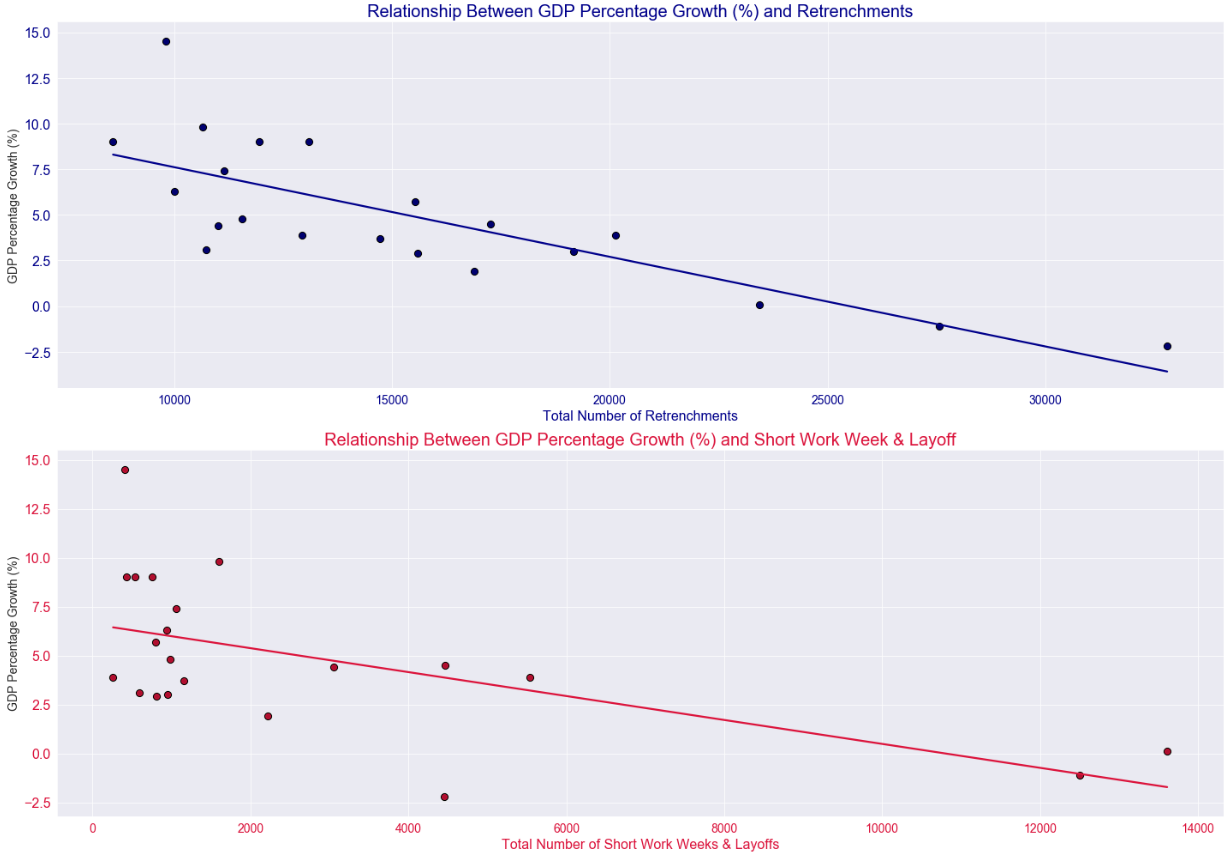
	Number				
Year	**2000**	**2007**	**2010**	**2014**	**2018**
Type					
Short Work Week and Temporary Layoffs (Total)	540	430	410	260	590

Write Python code that uses Python Visualisation Package (Matplotlib/Seaborn..) to produce useful data visualizations that explain the data.

Data Visualisation 5: Scatterplot (Seaborn and Matplotlib)

In [8]:

```
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib.ticker as ticker

#set seaborn background to darkgrid
sns.set_style("darkgrid")

#read from gdp growth file
gdp_growth = pd.read_csv("./Datasets/contribution-to-growth-in-gross-domestic-pro

#renaming gdp_growth column names
gdp_growth.rename(columns=
    {"year":"Year",
     "level_1":"GDP Growth in Chained Dollars (Base Year = 2015)",
     "value":"Percentage Growth (%)"}
    ,inplace=True)

#merging retrenchments, short work week and temporary layoff dataframe with gdp g
df_everything = pd.merge(df_all,gdp_growth,how="inner",on="Year")

#convert column for total retrenchments to type and Column values to row values
df_melt_retrenchment = df_everything.melt(value_vars=df_all[["Total Retrenchments

#convert column for total short work week and temporary layoff to type and Column
df_melt_short_work_week_and_layoff = df_everything.melt(value_vars=df_all[["Short

#declare fig and axes to plot
fig,ax = plt.subplots(2,1,figsize=(20,14))

#First Graph for Retrenchments

#declare ax0 to plot on the first subplot
ax0 = plt.subplot(211)

#plot using seaborn scatterplot method
sns.scatterplot(x="Number",y="Percentage Growth (%)", data=df_melt_retrenchment,

#plot using seaborn regplot to plot best fit line
sns.regplot(x="Number",y="Percentage Growth (%)", data=df_melt_retrenchment, colo

#set title for first graph plotting
ax0.set_title("Relationship Between GDP Percentage Growth (%) and Retrenchments",

#Second Graph for total short work week and temporary layoff

#declare ax1 to plot on the second subplot
ax1 = plt.subplot(212)

#plot using seaborn scatterplot method
sns.scatterplot(x="Number",y="Percentage Growth (%)", data=df_melt_short_work_wee

#plot using seaborn regplot to plot best fit line
sns.regplot(x="Number",y="Percentage Growth (%)", data=df_melt_short_work_week_an

#set title for second graph plotting
ax1.set_title("Relationship Between GDP Percentage Growth (%) and Short Work Week
#set fontsize, color for x labels, y labels, x ticks and y ticks
ax0.tick_params(axis="x", labelsize=14, colors='darkblue'), ax1.tick_params(axis=

ax0.tick_params(axis="y", labelsize=16, colors='darkblue'), ax1.tick_params(axis=

ax0.set_xlabel('Total Number of Retrenchments', fontsize=16, color='darkblue'), a

ax0.set_ylabel('GDP Percentage Growth (%)', fontsize=14), ax1.set_ylabel("GDP Per

#to display both graph nicely
fig.tight_layout()

#display graph
plt.show()
```



For the chosen datasets, explain the nature of that dataset (i.e. what is in that dataset) or any pecularities about it you wish to highlight and explain the process you went through to analyse that dataset, . Where possible, you should specifically mention how you used the Pandas, Matplotlib, Seaborn functions to achieve a certain outcome e.g. to transform the data or to produce a certain visualization:

Pecularities to highlight:

The total number of short work week & layoffs is much lesser than total retrenchments, with many data points hovering between values of 0-2000, which could have resulted in relative weak relationship between both variables. One possible reason for this is because Employers tend to only apply such measures during extreme circumstances, compared to retrenchments.

Process of using Pandas, Matplotlib or Seaborn functions to transform the data:

Using Pandas to transform

Firstly, I read the excel file from retrenchments and short work week and layoffs and store it in two variables. Then, I merge both dataframes into one and store it as df_all. After that, I display the shape, index and column names of the df_all dataset so that I am able to rename the column names. Moving on, I used the head and tail method to display the first and last five rows. To gather summary information of the data, I used the info method. The, I display Statistical Information of Retrenchments, Short Work Week and Temporary Layoffs Dataset using the describe method.

In addition to that, I am also interested to know the top 5 years where there are high/low retrenchment rates, high/low total number of short work week and layoff. In order to display this information, I first converted the column for total retrenchments to type and Column values to row values. In order to display the top 5 years where there are high retrenchment rates, I sorted the values by type, in ascending order, using the head method and store it as a variable. Then, I displayed the top 5 retrenchments using the pivot function. Similarly, the method to display top 5 years with lowest retrenchment rates is the same except for using the tail method. The way to display the top 5 years for total short work week and layoff is also the same as the retrenchment rates.

Using Matplotlib or Seaborn functions to transform the data:

Data Visualisation 5: Scatterplot

Firstly, I set style of Seaborn Graph background to dark and read from the gdp growth excel file. After that, I rename the column names for gdp growth and dropped level 1 columns which identifies each row as the GDP Growth Rate. Then, I merged retrenchments, short work week and temporary layoff dataframe with gdp growth on year into one dataframe called df_all. In order to display total retrenchments and total short work week and layoff using seaborn, I have to convert column for total retrenchments to type and Column values to row values as well as total short work week and layoff.

Then, I declare figure and axes object to plot. Then, I declare ax0 on the first plot and plot the graph using seaborn scatterplot and regplot (best fit line) using df_melt_retrenchment dataframe. Then, I set the title for the first subplot. Similarly, the process to plot the Second Graph for total short work week and temporary layoff is the same except for specifying ax1 on the second plot and using a df_total_short_work_week and layoff dataframe. In order to differentiate between both graphs, I specify the colors for retrenchments and total_short_week and layoffs to crimson and darkblue respectively for xticks, yticks, xlabels and title. Then, I call fig tight layout so that both graph will not interfere one another and plt show to display the graph.

For each dataset, highlight the insights you have gained from analysing the data and any conclusions or recommendations you want to make as a result of the analysis:

Data Visualisation 5: Scatterplot

Since the scatterplot is used to distinguish the relationship between two variables, the scatterplot is the last figure plotted for data visualisation to further justify that the insights gained from such analysing the dataset is accurate.

From the graph, I are able to conclude that the hypothesis of a strong correlation between GDP Growth, retrenchments and short work week & layoffs is correct. This is evident as there is strong linear relationship between GDP Percentage Growth and retrenchments, as well as short work week and layoffs. When there are high GDP Growth Rates, the total number of retrenchments is one of the lowest. When there are low GDP Growth Rates, the total number of retrenchments/short work week and layoff is one of the highest. As such, this shows that Singapore is heavily influenced by Global Events as it has a strong impact on the number of retrenchments, short work week and layoff.

Key Takeaways from Data Analysis

- Singapore's Prosperous Economy is largely due to its International financial trade center, which makes it attractive for foreign investments in Singapore.
- Given the image that Singapore portrays to the world, the success of Singapore's Economy is largely intertwined with the Global Economy.
- As such, Singapore is very vulnerable to Global Economic Crisis as it relies heavily on global trade demands.
- Looking forward, Singapore needs to constantly reinvent itself in order to survive for challenges in the decades to come.