

DASHBOARD

NAME : SANDHYA GOYAL

GROUP : CS8

ROLL NO : 102017182

DATASET

VARIABLES	DESCRIPTION
Country or Area	Three Countries – <u>Austrilia</u> , <u>Canada</u> , <u>USA</u>
Year	Year range 1988-2016
Commodity Code	Item code <u>unqie</u> for all
Commodity	Type of item
Flow	Indicate flow of trade- Export Import Re-Import Re-Export
Trade USD	Trade value is US Dollars

DATASET

Weight	Total weight of traded items in Kgs
Quantity Name	Quantity name of traded items
Quantity	Total quantity of the traded items
Category	12 unique categories

```
library(dplyr)
trade <- read.csv("Raw Data.csv")

view(trade)
```

	country_or_area	year	comm_code	commodity	flow	trade_usd	weight_kg	quantity_name	quantity
1	Australia	2016	10111	Horses, live pure-bred breeding	Export	128577553	900450	Number of items	1882
2	Australia	2016	10119	Horses, live except pure-bred breeding	Re-Import	4928989	47240	Number of items	104
3	Australia	2016	10119	Horses, live except pure-bred breeding	Export	11812782	153587	Number of items	276
4	Australia	2016	10119	Horses, live except pure-bred breeding	Import	90430302	1082493	Number of items	2073
5	Australia	2016	10120	Asses, mules and hinnies, live	Export	58473	5805	Number of items	416
6	Australia	2016	10120	Asses, mules and hinnies, live	Import	5221	1000	Number of items	2
7	Australia	2016	10210	Bovine animals, live pure-bred breeding	Export	138293019	35071238	Number of items	116088
8	Australia	2016	10290	Bovine animals, live, except pure-bred breeding	Export	947455270	356789795	Number of items	1036209
9	Australia	2016	10310	Swine, live pure-bred breeding	Export	39434	664	Number of items	24
0	Australia	2016	10410	Sheep, live	Export	162261760	87199485	Number of items	1870412
1	Australia	2016	10410	Sheep, live	Import	137397	16384	Number of items	170
2	Australia	2016	10420	Goats, live	Export	5951870	2116875	Number of items	54703
3	Australia	2016	10511	Fowls, live domestic < 185 grams	Export	1532654	14040	Number of items	215968
4	Australia	2016	10600	Animals, live, except farm animals	Re-Import	5183	84	Number of items	3
5	Australia	2016	10600	Animals, live, except farm animals	Export	2853075	70647	Number of items	43944

```
#make a copy and work on the copy
business <- trade

#Modify names of columns for clarity
names(business) <- c("Country", "Year", "Commoditycode", "Commodity", "Flow",
                     "Dollars", "Weight", "Quantityname", "Quantity", "Category")

#structure of data
str(business)
```

```
> business <- trade
> #Modify names of columns for clarity
> names(business) <- c("Country", "Year", "Commoditycode", "Commodity", "Flow",
+                      "Dollars", "Weight", "Quantityname", "Quantity", "Category")
> #structure of data
> str(business)
'data.frame':   59090 obs. of  10 variables:
 $ Country      : chr  "Australia" "Australia" "Australia" "Australia" ...
 $ Year         : int   2016 2016 2016 2016 2016 2016 2016 2016 2016 2016 ...
 $ Commoditycode: int   10111 10119 10119 10119 10120 10120 10210 10290 10310 10410 ...
 $ Commodity    : chr   "Horses, live pure-bred breeding" "Horses, live except pure-bred breeding" "Horses, live except pure-bred breeding" "Horses, live except pure-bred breeding" ...
 $ Flow         : chr   "Export" "Re-Import" "Export" "Import" ...
 $ Dollars      : num   1.29e+08 4.93e+06 1.18e+07 9.04e+07 5.85e+04 ...
 $ Weight       : num   900450 47240 153587 1082493 5805 ...
 $ Quantityname : chr   "Number of items" "Number of items" "Number of items" "Number of items" ...
 $ Quantity     : num   1882 104 276 2073 416 ...
 $ Category     : chr   "01_live_animals" "01_live_animals" "01_live_animals" "01_live_animals" ...
```

```
#summary of data
summary(business)
```

```
> #summary of data
> summary(business) #Missing values present in Weight and Quantity
  Country      Year  Commoditycode  Commodity
Length:59090   Min.   :1988      Min.   : 10111   Length:59090
Class :character 1st Qu.:1996      1st Qu.: 30559   Class :character
Mode  :character Median :2003      Median : 70690   Mode  :character
                Mean  :2003      Mean  : 64722
                3rd Qu.:2010      3rd Qu.: 90620
                Max.   :2016      Max.   :121490

  Flow      Dollars      weight      Quantityname
Length:59090   Min.   :1.000e+00   Min.   :0.000e+00   Length:59090
Class :character 1st Qu.:6.785e+04   1st Qu.:1.829e+04   Class :character
Mode  :character Median :8.304e+05   Median :2.747e+05   Mode  :character
                Mean  :3.107e+07   Mean  :8.739e+07
                3rd Qu.:7.308e+06   3rd Qu.:3.174e+06
                Max.   :1.373e+10   Max.   :6.140e+11
                NA's   :715

  Quantity      Category
Min.   :0.000e+00   Length:59090
1st Qu.:1.901e+04   Class :character
Median :2.854e+05   Mode  :character
```

```
mean   :8.806e+07
3rd Qu.:3.317e+06
Max.   :6.140e+11
NA's   :705
```

```
#Approach 1: Eliminate all missing values
trader <- na.omit(trade)
```

Approach 2: Understand why there are missing values

```
#create new columns for missing weights and quantities
#When weight is zero and Quantity is also zero, output is 1 else 0
business$zerowQ <- ifelse(business$Weight < 1 & business$Quantity < 1, "1", "0")
table(business$zerowQ)# there are 847 rows that account to this scenario
business$zerowQ <-NULL

#When Weight is zero and Quantity is has NA, output is 1 else 0
#we can discard this condition
business$zerow_naQ <- ifelse(business$Weight < 1 & business$Quantity == NA, "1", "0")

table(business$zerow_naQ)

business$zerow_naQ <- NULL
```

```
#When Weight is zero, or NAs in Weight, or Quantity is zero, or NAs in Quantity - assign 1
business$Missing <- ifelse(business$Weight < 1 | is.na(business$Weight)
                           | business$Quantity < 1 | is.na(business$Quantity), "1", "0")
table(business$Missing)
```

> table(business\$zerowQ)# there are 847 rows that account to this scenario

```
      0      1
57855  847
> business$zerowQ <-NULL
> #When Weight is zero and Quantity is has NA, output is 1 else 0
> #we can discard this condition
> business$zerow_naQ <- ifelse(business$Weight < 1 & business$Quantity == NA, "1", "0")
> table(business$zerow_naQ)

      0
56590
> business$zerow_naQ <- NULL
> #When Weight is zero, or NAs in Weight, or Quantity is zero, or NAs in Quantity - assign 1
> business$Missing <- ifelse(business$Weight < 1 | is.na(business$Weight)
                             | business$Quantity < 1 | is.na(business$Quantity), "1", "0")
> table(business$Missing)

      0      1
56771  2819
```

```
#there are 2819 cases, where Weight and Quantity have either zero value or NA
#This column will help us identify the Trade value(in USD) incurred for missing info.
#understand the consequences of missing info by country
missedinfo_austr <- filter(business, Country == "Australia", Missing == 1)
nrow(missedinfo_austr) # 446 rows have missing info on weight and quantity columns
sum(missedinfo_austr$Dollars)# 13.5 billion US Dollars (13586563312)
```

- #there are 2819 cases, where weight and quantity have either zero value or NA
- #This column will help us identify the Trade value(in USD) incurred for missing info.
- #understand the consequences of missing info by country
- missedinfo_australia <- filter(business, Country == "Australia", Missing == 1)
- nrow(missedinfo_australia) # 446 rows have missing info on Weight and Quantity columns
- [1] 446
- sum(missedinfo_australia\$Dollars) # 13.5 billion US Dollars (13586563312)
- [1] 13586563312

```
missedinfo_can <- filter(business, Country == "Canada", Missing == 1)
nrow(missedinfo_can) # 1429 rows have missing info on Weight and Quantity
sum(missedinfo_can$Dollars) #32.5 billion US Dollars (32510198277)
```

```
> missedinfo_can <- filter(business, Country == "Canada", Missing == 1)
> nrow(missedinfo_can) # 1429 rows have missing info on Weight and Quantity
[1] 1429
> sum(missedinfo_can$Dollars) #32.5 billion US Dollars (32510198277)
[1] 32510198277
```

```
missedinfo_usa <- filter(business, Country == "USA", Missing == 1)
nrow(missedinfo_usa) # 994 rows have missing info on Weight and Quantity
sum(missedinfo_usa$Dollars) #36.9 billion US Dollars (36952700815)
```

```
> missedinfo_usa <- filter(business, Country == "USA", Missing == 1)
> nrow(missedinfo_usa) # 994 rows have missing info on Weight and Quantity
[1] 994
> sum(missedinfo_usa$Dollars) #36.9 billion US Dollars (36952700815)
[1] 36952700815
```

```
#understand the consequences of missing info by Flow
#AUSTRALIA
aus_export <- filter(missedinfo_australia, Flow == "Export")
nrow(aus_export) # 52 rows have missing info on Weight and Quantity
sum(aus_export$Dollars) #120325935471 12 Billion US Dollars from rows containing missing info
```

```
> #understand the consequences of missing info by Flow
> #AUSTRALIA
> aus_export <- filter(missedinfo_aus, Flow == "Export")
> nrow(aus_export) # 52 rows have missing info on Weight and Quantity
[1] 52
> sum(aus_export$Dollars) #(12032593547) 12 Billion US Dollars from rows containing missing info
[1] 12032593547
```

```
aus_import <- filter(missedinfo_aus, Flow == "Import")
nrow(aus_import) # 383 rows have missing info on Weight and Quantity
sum(aus_import$Dollars) #(1552825617) 1.5 Billion US Dollars from rows containing missing info
```

```
> aus_import <- filter(missedinfo_aus, Flow == "Import")
> nrow(aus_import) # 383 rows have missing info on Weight and Quantity
[1] 383
> sum(aus_import$Dollars) #(1552825617) 1.5 Billion US Dollars from rows containing missing info
[1] 1552825617
```

```
aus_reimport <- filter(missedinfo_aus, Flow == "Re-Import")
nrow(aus_reimport) # 8 rows have missing info on Weight and Quantity
sum(aus_reimport$Dollars) #(938823) 938,823 US Dollars from rows containing missing info

aus_reexport <- filter(missedinfo_aus, Flow == "Re-Export")
nrow(aus_reexport) # 3 rows have missing info on Weight and Quantity
sum(aus_reexport$Dollars) #(205325) 205,325 US Dollars from rows containing missing info
```

```
> nrow(aus_reimport) # 8 rows have missing info on Weight and Quantity
[1] 8
> sum(aus_reimport$Dollars) #(938823) 938,823 US Dollars from rows containing missing info
[1] 938823
> aus_reexport <- filter(missedinfo_aus, Flow == "Re-Export")
> nrow(aus_reexport) # 3 rows have missing info on Weight and Quantity
[1] 3
> sum(aus_reexport$Dollars) #(205325) 205,325 US Dollars from rows containing missing info
[1] 205325
```

```
can_export <- filter(missedinfo_can, Flow == "Export")
nrow(can_export) # 637 rows have missing info on Weight and Quantity
sum(can_export$Dollars) #(26578595902) 26.5 Billion US Dollars from rows containing missing info

can_import <- filter(missedinfo_can, Flow == "Import")
nrow(can_import) # 582 rows have missing info on Weight and Quantity
sum(can_import$Dollars) #(5838350510) 5.8 Billion US Dollars from rows containing missing info

can_reimport <- filter(missedinfo_can, Flow == "Re-Import")
nrow(can_reimport) # 100 rows have missing info on Weight and Quantity
sum(can_reimport$Dollars) #(3676340) 3.6 Million US Dollars from rows containing missing info

can_reexport <- filter(missedinfo_can, Flow == "Re-Export")
nrow(can_reexport) # 110 rows have missing info on Weight and Quantity
sum(can_reexport$Dollars) #(89575525) 89 Million US Dollars from rows containing missing info
```



```

> #CANADA
> can_export <- filter(missedinfo_can, Flow == "Export")
> nrow(can_export) # 637 rows have missing info on Weight and Quantity
[1] 637
> sum(can_import$Dollars) #(5838350510) 5.8 Billion US Dollars from rows containing missing info
[1] 5838350510
> can_reimport <- filter(missedinfo_can, Flow == "Re-Import")
> nrow(can_reimport) # 100 rows have missing info on Weight and Quantity
[1] 100
> sum(can_reimport$Dollars) #(3676340) 3.6 Million US Dollars from rows containing missing info
[1] 3676340
> can_reexport <- filter(missedinfo_can, Flow == "Re-Export")
> nrow(can_reexport) # 110 rows have missing info on Weight and Quantity
[1] 110
> sum(can_reexport$Dollars) #(89575525) 89 Million US Dollars from rows containing missing info
[1] 89575525
>

```

```

#USA
usa_export <- filter(missedinfo_usa, Flow == "Export")
nrow(usa_export) # 385 rows have missing info on Weight and Quantity
sum(usa_export$Dollars) #(17946793912) 17.9 Billion US Dollars from rows containing missing info

usa_import <- filter(missedinfo_usa, Flow == "Import")
nrow(usa_import) # 305 rows have missing info on Weight and Quantity
sum(usa_import$Dollars) #(18418691692) 18.4 Billion US Dollars from rows containing missing info

usa_reimport <- filter(missedinfo_usa, Flow == "Re-Import")
nrow(usa_reimport) # 0 rows have missing info on Weight and Quantity
sum(usa_reimport$Dollars) #0 US Dollars from rows containing missing info

usa_reexport <- filter(missedinfo_usa, Flow == "Re-Export")
nrow(usa_reexport) # 254 rows have missing info on Weight and Quantity
sum(usa_reexport$Dollars) #(587215211) 5.8 Billion US Dollars from rows containing missing info

```

```

#USA
> usa_export <- filter(missedinfo_usa, Flow == "Export")
> nrow(usa_export) # 385 rows have missing info on Weight and Quantity
[1] 385
> sum(usa_export$Dollars) #(17946793912) 17.9 Billion US Dollars from rows containing missing info
[1] 17946793912
> usa_import <- filter(missedinfo_usa, Flow == "Import")
> nrow(usa_import) # 305 rows have missing info on Weight and Quantity
[1] 305
> sum(usa_import$Dollars) #(18418691692) 18.4 Billion US Dollars from rows containing missing info
[1] 18418691692
> usa_reimport <- filter(missedinfo_usa, Flow == "Re-Import")
> nrow(usa_reimport) # 0 rows have missing info on Weight and Quantity
[1] 0
> sum(usa_reimport$Dollars) #0 US Dollars from rows containing missing info
[1] 0
> usa_reexport <- filter(missedinfo_usa, Flow == "Re-Export")
> nrow(usa_reexport) # 254 rows have missing info on Weight and Quantity
[1] 254
> sum(usa_reexport$Dollars) #(587215211) 5.8 Billion US Dollars from rows containing missing info
[1] 587215211

```

```

#UNDERSTAND THE TOTAL USD for each country from business dataset
#AUSTRALIA
totalusd_aus <- filter(business, Country == "Australia", Dollars >= 1)
nrow(totalusd_aus) #Australia occupies 24,921 rows in the dataset
sum(totalusd_aus$Dollars) #Total US dollars account to 458 billion
#excluding rows which have missing info
usd_aus <- filter(totalusd_aus, Missing == 0)
nrow(usd_aus) #24475 rows with complete info
sum(usd_aus$Dollars)# 445 Billion US Dollars (Difference)
options(scipen = 999)
#filter above by Flow
expo_aus <- filter(usd_aus, Flow == "Export")
nrow(expo_aus) #11133 rows with complete info
sum(expo_aus$Dollars)# 382 Billion US Dollars
imp_aus <- filter(usd_aus, Flow == "Import")
nrow(imp_aus) #9964 rows with complete info
sum(imp_aus$Dollars)# 62.6 Billion US Dollars
reimp_aus <- filter(usd_aus, Flow == "Re-Import")
nrow(reimp_aus) #1335 rows with complete info
sum(reimp_aus$Dollars)# 0.27 Billion US Dollars
reexpo_aus <- filter(usd_aus, Flow == "Re-Export")
nrow(reexpo_aus) #2043 rows with complete info
sum(reexpo_aus$Dollars)# 0.19 Billion US Dollars

```

```

#CANADA
totalusd_can <- filter(business, Country == "Canada", Dollars >= 1)
nrow(totalusd_can) #Canada occupies 29,932 rows in the dataset
sum(totalusd_can$Dollars) #Total US dollars account to 775 billion

#excluding rows which have missing info
usd_can <- filter(totalusd_can, Missing == 0)
nrow(usd_can) #28,503 rows with complete info
sum(usd_can$Dollars)# 742 Billion US Dollars (Difference)
#filter above by Flow
expo_can <- filter(usd_can, Flow == "Export")
nrow(expo_can) # 10592 rows with complete info
sum(expo_can$Dollars)# 488 Billion US Dollars

imp_can <- filter(usd_can, Flow == "Import")
nrow(imp_can) #11248 rows with complete info
sum(imp_can$Dollars)# 249 Billion US Dollars

reimp_can <- filter(usd_can, Flow == "Re-Import")
nrow(reimp_can) #3307 rows with complete info
sum(reimp_can$Dollars)# 0.5 Billion US Dollars

reexpo_can <- filter(usd_can, Flow == "Re-Export")
nrow(reexpo_can) #3356 rows with complete info
sum(reexpo_can$Dollars)# 4 Billion US Dollars

```



```

#USA
totalusd_usa <- filter(business, Country == "USA", Dollars >= 1)
nrow(totalusd_usa) #USA occupies 4237 rows in the dataset
sum(totalusd_usa$Dollars) #Total US dollars account to 601 billion

#excluding rows which have missing info
usd_usa <- filter(totalusd_usa, Missing == 0)
nrow(usd_usa) #3293 rows with complete info
sum(usd_usa$Dollars)# 564 Billion US Dollars (Difference)
#filter above by Flow
expo_usa <- filter(usd_usa, Flow == "Export")
nrow(expo_usa) #1144 rows with complete info
sum(expo_usa$Dollars)# 429 Billion US Dollars

imp_usa <- filter(usd_usa, Flow == "Import")
nrow(imp_usa) #1210 rows with complete info
sum(imp_usa$Dollars)# 133 Billion US Dollars

reimp_usa <- filter(usd_usa, Flow == "Re-Import")
nrow(reimp_usa) #0 rows with complete info
sum(reimp_usa$Dollars)# 0 US Dollars

reexpo_usa <- filter(usd_usa, Flow == "Re-Export")
nrow(reexpo_usa) #939 rows with complete info
sum(reexpo_usa$Dollars)# 1.9 Billion US Dollars

```

#INSIGHTS FROM DATA

#USA has not re-imported any commodity from 1988 - 2016

#There have been import bills and export payouts for data having no info of commodities weight and quantity.

#USA leads with 36.9 billion US Dollars for data having missing info

#Canada follows USA with 32.5 billion US Dollars

#While Australia generated 13.5 billion US Dollars (Missing info highest in Export and lowest in Re export Flow)

#Total Trade in US Dollars in descending order (business dataset)

#CANADA - 775 BILLION

#USA - 601 BILLION

#AUS - 458 BILLION

#Excluding rows with missing info: Ranking in Largest Trade Share

#CANADA - 742 BILLION

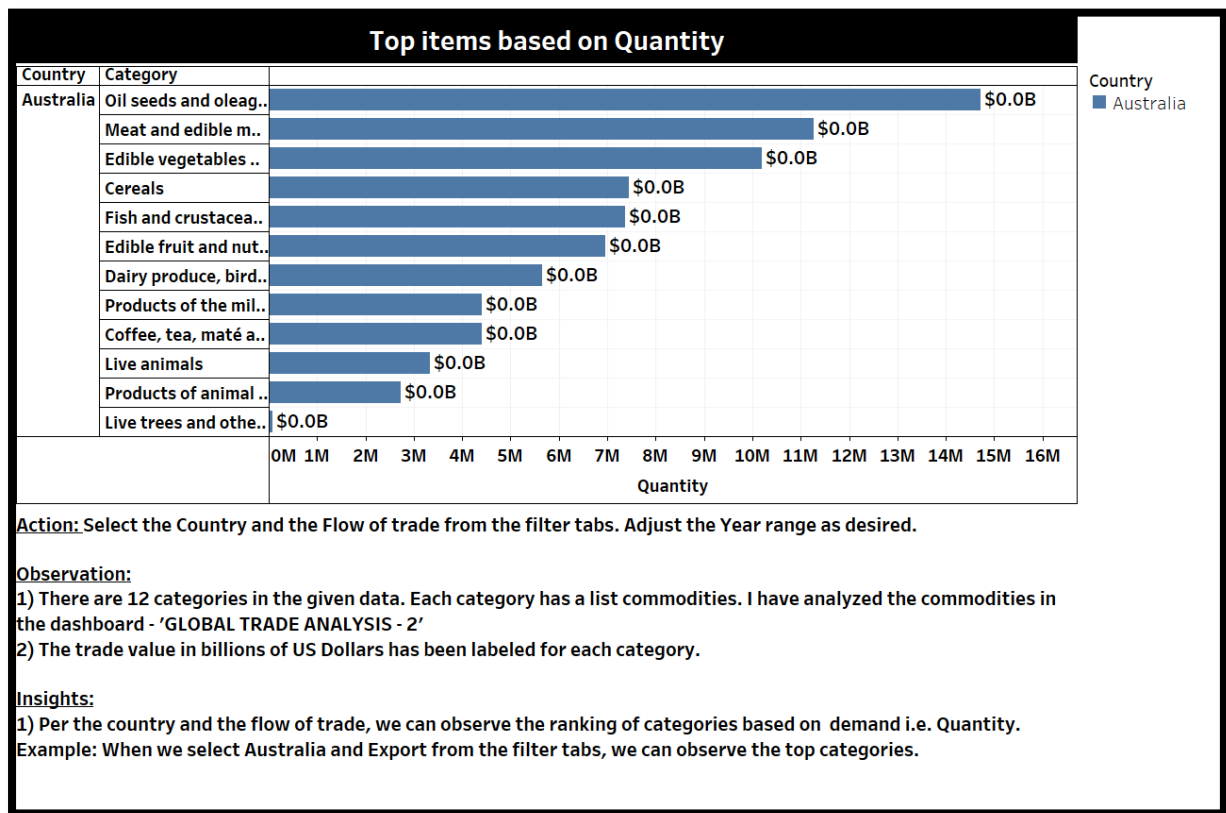
#USA - 564 BILLION

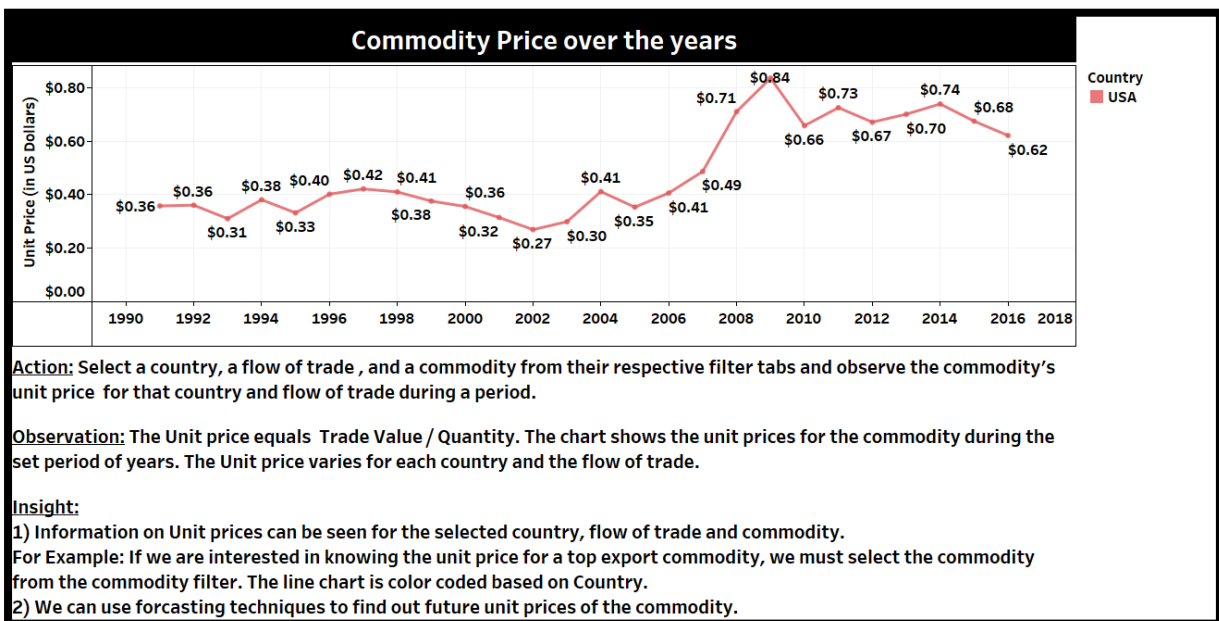
#AUS - 445 BILLION

INSIGHTS FROM THE GRAPHS:

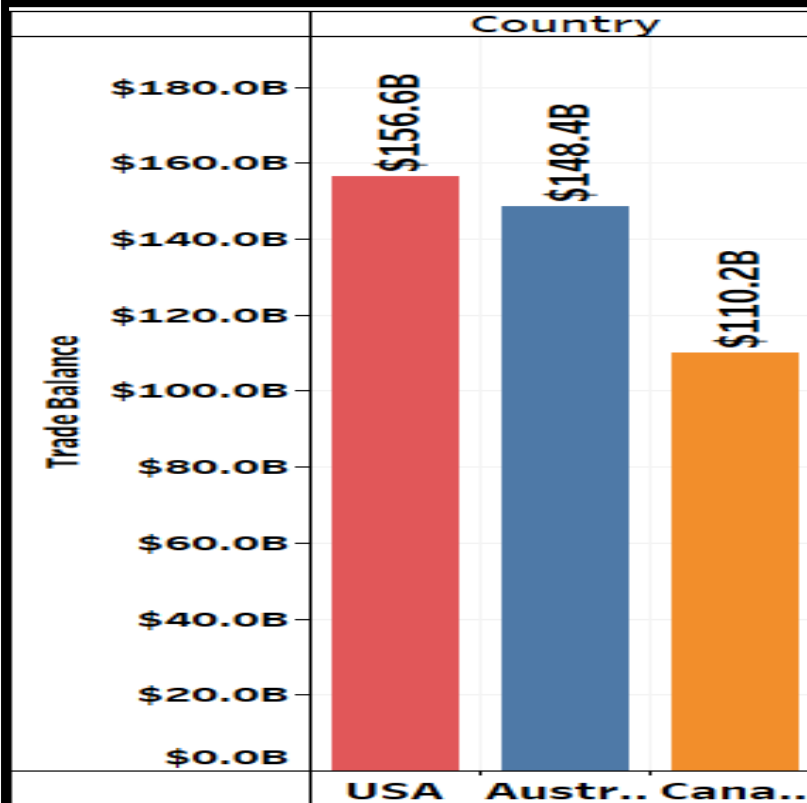
1. We can get data country wise . Just select country for which you want to visualize the import and export graphs
2. We can also set a range for years by moving the slider (Ex 2012-2016)
3. Different items import export quantity
4. Comparison between different countries on basis of quantity (filter is import , export, re export , re import)
5. Compare on the basis of trade balance analysis

GRAPHS:





Trade Balance Analysis



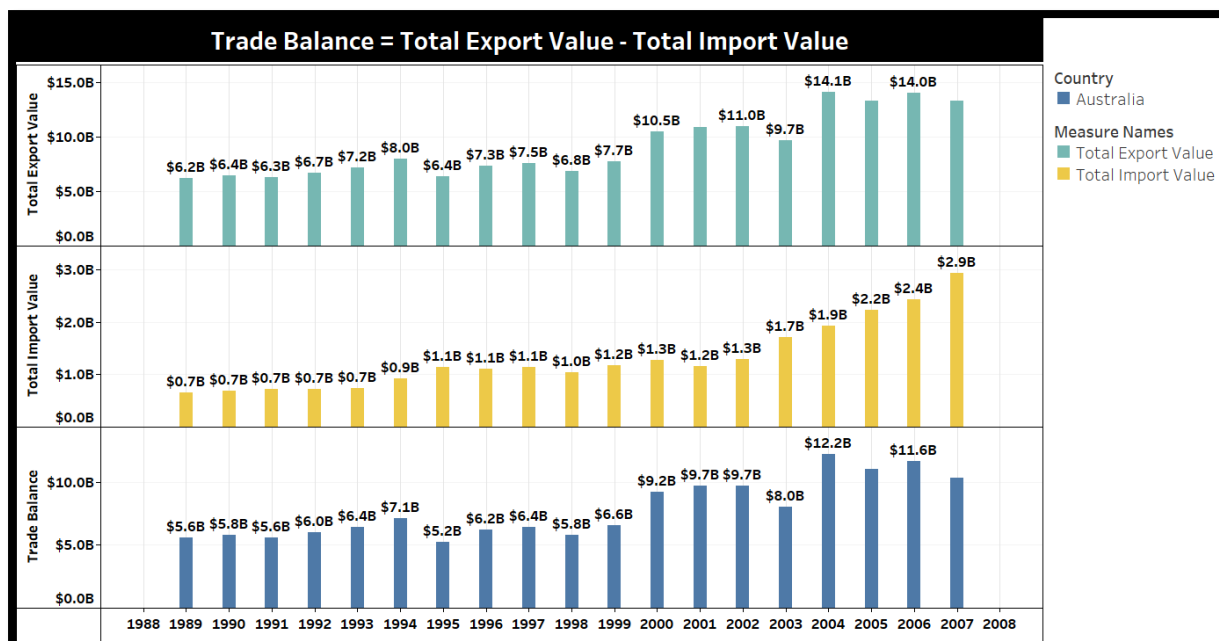
Country
 USA
 Australia
 Canada

Action: Select the Country and adjust the range of Year to observe the overall Trade Balance for those years.

Observation: We can observe the ranking of Countries by their Trade Balance.

Insight:

- 1) The Ranking is in this order - Australia > USA > Canada for the years 1988-2016. While, for the years 2012 - 2016 the ranking is as follows - Australia > Canada > USA. Considering 2016 alone, the ranking is the same as the previous setting (2012 - 2016).
- 2) Trade Surplus countries may not be suitable for Import business, as countries tend to levy higher import duty on commodities. This is done so that domestic goods are consumed more than foreign goods by their people.



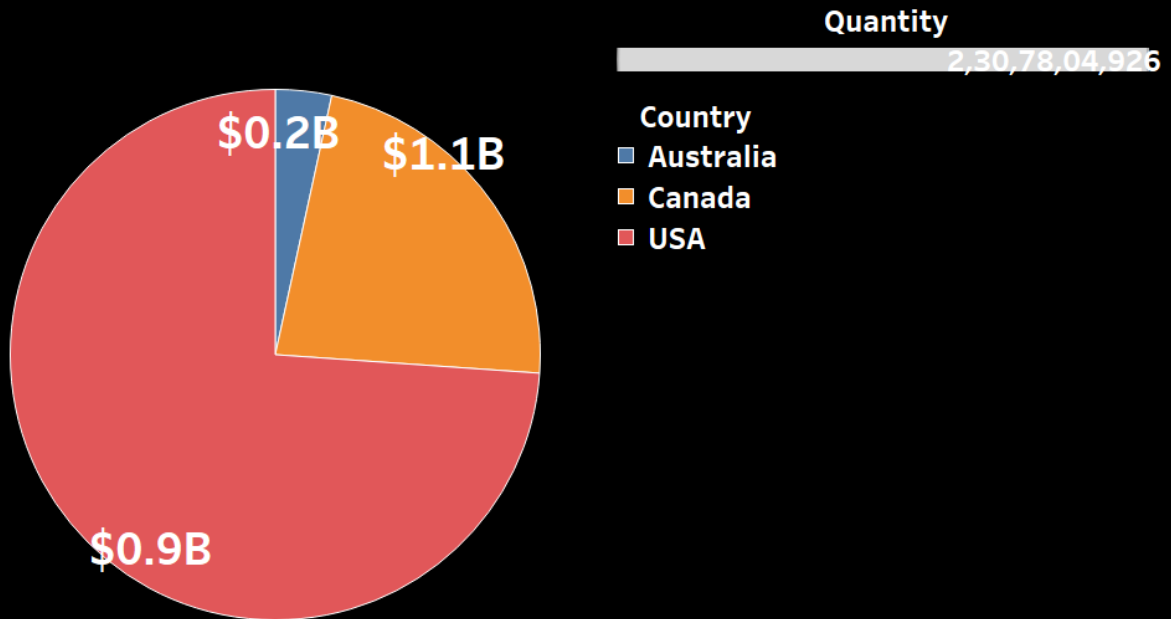
Action: Select the Country and the Flow of trade from the filter tabs. Adjust the Year range as desired.

Observation: The Trade balance has been calculated by subtracting the total import value from the total export value. It measures the country's net income earned on international assets. If the Trade Balance is positive, then it is called Trade Surplus, while a negative value is called Trade Deficit. Most countries regard Trade Deficit as unfavorable. This indicates the country imports commodities more than it exports them. However, a trade surplus may not be in the country's best interests.

Insight:

1) The Trade Balance has been positive for the three countries throughout the years 1988 - 2016, indicating Trade Surplus.

Market Size based on Quantity



Sum of Trade in USD. Color shows details about Country. Size shows sum of Quantity. The marks are labeled by sum of Trade in USD. The data is filtered on Flow and Year. The Flow filter keeps Re-Export. The Year filter ranges from 1989 to 2007.

Word Cloud

Country

- USA

Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage

Cereals

Live animals

Products of animal origin, not elsewhere specified or included

Category. Color shows details about Country. Size shows sum of Quantity. The data is filtered on Flow and Year. The Flow filter keeps Export. The Year filter ranges from 2010 to 2016. The view is filtered on Country, which keeps USA.

DASHBOARD

