☢️

EVE-War Eng

EVE WAR_EDA (Explored number of alliance in the war data)

⚠️ Goal:

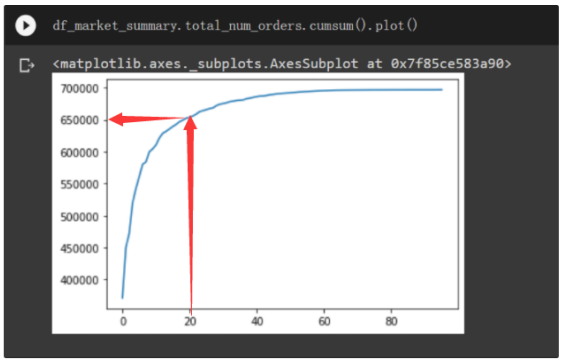
- Introduce total_value_destroyed, total_value_dropped
  - Properly Connect **Market** table to **Item** table
    - Idea Market Keys: Date, Item_Type, Region
    - Idea Item Keys: Item_Type, Killmail_id
  - Aggregate **Item** table to make <Item_Type, Killmail_id> composite primary key
- **Introduce the following attacker related features:**
  - Major attacker alliance
  - Character_Final_Blow
  - Attacker_Avg_Security
  - Num_Attackers
- Aggregate Across War_ID to have the following features:
  - Total_Value_Destroyed_By_Alliance_A
  - Total_Value_Destroyed_By_Alliance_B
  - Total_Value_Dropped_By_Alliance_A
  - Total_Value_Dropped_By_Alliance_B
  - Total_Damage_Taken_By_Alliance_A
  - Total_Damage_Taken_By_Alliance_B

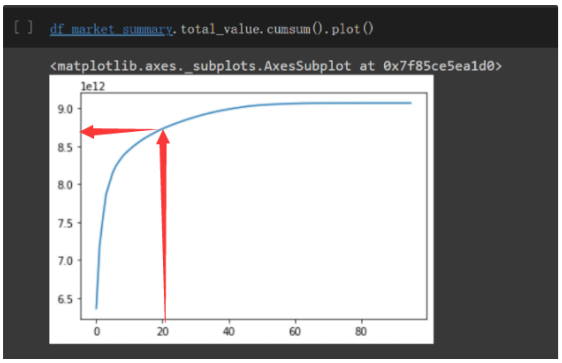
1. Aggregate **Item** table to make <Item_Type, Killmail_id> composite primary key

```
create table market.item_rdtj
as
(
select killmail_id, kdt, item_type_id, sum(quantity_destroyed) as quantity_destroyed_tot, sum(quantity_dropped) as quantity_dropped_tot
from (
  select w.killmail_id,  date(w.killmail_time) as kdt, item_type_id,quantity_destroyed,quantity_dropped
  from market.War w left join market.warItems i
  on w.killmail_id = i.killmail_id) as t1
group by  killmail_id,kdt, item_type_id
)
```

2. Define & Explore popularity

```
df_market_daily = market.withColumn('value',col('average')*col('order_count')).groupBy("date", "region_id").\
                                agg(countDistinct('type_id').alias('diversity'),
                                          sum('value').alias('total_value') ,
                                          sum('volume').alias('total_volume') ,
                                          sum('order_count').alias('total_num_orders'))
df_market_summary = df_market_daily_pd.groupby('region_id').mean().sort_values(['total_value','total_num_orders','diversity'],ascending
```





3. Prepare Market Table by choosing the TOP 20 Popular regions, using the average among such subset to define market price

```
create table market.market_rdtj
as
(select date,type_id, avg(average) as market_price_avg
from (select *
from market.market_history h right join
    (select *
  from market.Market_Region_Summary
  limit 20) as t1 on
  h.region_id = t1.region_id) as t2
group by date,type_id
order by date)
```

4. Join the 2 tables together into **Item_Market** for future aggregation

```
create table market.item_market
as
(select killmail_id,date,type_id,quantity_destroyed_tot*market_price_avg as value_destroyed_tot, quantity_dropped_tot * market_price_av
from `proud-apogee-348022.market.item_rdtj` I left join
  `proud-apogee-348022.market.market_rdtj` M
  on I.item_type_id = M.type_id and I.kdt = M.date)
```

5. Select the wars where only 2 alliances participated

```
create table market.WARS_DUO
AS
(
SELECT W.*, t1.Num_Alliance
FROM
    (SELECT war_id,count(DISTINCT victim_alliance_id ) as Num_Alliance
FROM `proud-apogee-348022.market.War` W
GROUP BY war_id) AS t1
    LEFT JOIN `proud-apogee-348022.market.War` W on t1.war_id = W.war_id
  where t1.Num_Alliance = 2
)
```