

Santa Catarina Import-Export

March 6, 2020

1 Import and Export Data for Santa Catarina

We want to answer the following questions:

- What are the top 3 most exported products by State for the years 2017, 2018 and 2019?
- What are the top 3 most imported products by State for the years 2017, 2018 and 2019?
- What are the top 3 most exported products in each month of 2019 by State?
- What is the percentage of total national exports by State in 2019?
- What is the percentage of total national imports by State in 2019?
- Prediction of value of top 3 exported products by month from Santa Catarina to each target Country.
- Prediction of value of top 3 imported products by month from Santa Catarina from each source Country.

1.1 Contents

1.2 Data Wrangling for Exports

First step to address our analysis will be to download the files, unzip them, clean them and finally organize the best way possible. I have decided to store the link to the government website and download it as we execute the notebook instead of uploading the file as Git LFS because those file locations **should not** change at all. Considering we are betting on the Brazilian government efficiency, we will store the MD5 of the used files “just in case” they ever change for any reason. One faster way to do this is simply download the zip files yourself and put them in the data folder.

```
EXP_COMPLETA.csv
EXP_COMPLETA.zip
IMP_COMPLETA.csv
IMP_COMPLETA.zip
NCM.csv
PAIS.csv
README.txt
VIA.csv
```

We start with exports. One important point we notice are the file sizes. One has around 1.5G unzipped, the other has 2.15G. Those may be a problem for python if you are using a computer with less than 8G RAM Memory.

Let's import the first one.

We will replace the names with more human friendly ones, and we notice that some of the data

First thing I learned was what FOB means! [https://www.ipea.gov.br/desafios/index.php?option=com_content&vi](https://www.ipea.gov.br/desafios/index.php?option=com_content&view=article&id=10786)

So, the column `VL_FOB` should be renamed to `Amount(USD)`. We also need to download the remaining tables with more human friendly definitions so we get our final Dataframe. Another important aspect to notice is that we will only need data from 2017 - 2019, so we will filter those in order to spare some memory space.

After some manipulations and merging of Dataframes, we finally have the ideal Dataframe with all information we need in a more human friendly naming system.

	Ano	Mes	Estado	Pais	Via \
0	2017	5	SC	Bolívia	RODOVIARIA
1	2017	5	PR	Bolívia	RODOVIARIA
2	2017	4	SC	Bolívia	RODOVIARIA
3	2017	8	SC	Bolívia	RODOVIARIA
4	2017	5	MG	Bolívia	RODOVIARIA
5	2017	8	PR	Bolívia	RODOVIARIA
6	2017	3	PR	Bolívia	RODOVIARIA
7	2017	2	PR	Bolívia	RODOVIARIA
8	2017	10	SC	Bolívia	RODOVIARIA
9	2017	10	MG	Bolívia	RODOVIARIA

	Produto	Kg	Amount (USD)
0	Combinações de refrigeradores e congeladores (...)	61327	312310
1	Combinações de refrigeradores e congeladores (...)	1092	3570
2	Combinações de refrigeradores e congeladores (...)	58772	281605
3	Combinações de refrigeradores e congeladores (...)	95000	491717
4	Combinações de refrigeradores e congeladores (...)	2970	16919
5	Combinações de refrigeradores e congeladores (...)	40324	228961
6	Combinações de refrigeradores e congeladores (...)	20213	66990
7	Combinações de refrigeradores e congeladores (...)	21078	75652
8	Combinações de refrigeradores e congeladores (...)	144439	662794
9	Combinações de refrigeradores e congeladores (...)	7138	42485

Next thing is to check for NAs, but there should not be many because all our merging was done using inner joins (we discarded information that was missing in either tables).

Ano	0
Mes	0
Estado	0
Pais	0
Via	0
Produto	0

```

Kg          0
Amount(USD) 0
dtype: int64

```

We now check if States, Countries and Months are correct.

```

0      SC
1      PR
4      MG
14     SP
271    MT
295    ND
686    AP
798    AL
800    ES
814    RS
815    MS
907    RJ
1308   PE
2161   BA
3213   CE
3374   RE
3852   RO
4293   GO
6535   RN
7833   AC
8134   RR
8160   AM
8389   PI
14301  PA
18735  DF
23753  SE
25935  PB
33009  MA
77219  TO
153119 MN
3566074 CB

```

```
Name: Estado, dtype: object
```

```

0      Bolívia
102    Argentina
113    Paraguai
176    Chile
184    Uruguai
...
1002410  Cocos (Keeling), Ilhas
1123306  Toquelau

```

```

1421174          Lebuan, Ilhas
1652762          Bouvet, Ilha
3390182  Marianas do Norte, Ilhas
Name: Pais, Length: 252, dtype: object

```

```

252

```

```

10      1
7       2
6       3
2       4
0       5
16      6
15      7
3       8
11      9
8       10
12      11
17      12
Name: Mes, dtype: int64

```

It seems we will have to remove 4 States that do not exist (RE, ND, MN and CB), the rest seems all right.

1.3 Analysing Exports

Next step in our analysis is to answer the first question: list the top 3 exports by estate in 2017, 2018 and 2019.

	Ano	Estado	Produto \
105237	2019	TO	Soja, mesmo triturada, exceto para semeadura
105174	2019	TO	Carnes desossadas de bovino, congeladas
105184	2019	TO	Milho em grão, exceto para semeadura
65064	2018	TO	Soja, mesmo triturada, exceto para semeadura
65021	2018	TO	Carnes desossadas de bovino, congeladas
65022	2018	TO	Carnes desossadas de bovino, frescas ou refrig...
29777	2017	TO	Soja, mesmo triturada, exceto para semeadura
29728	2017	TO	Carnes desossadas de bovino, congeladas
29739	2017	TO	Milho em grão, exceto para semeadura
105141	2019	SP	Óleos brutos de petróleo

	Amount (USD)
105237	772631871
105174	155101002
105184	78543244
65064	995302750
65021	88538892
65022	25932145

29777	755967278
29728	78877967
29739	53312114
105141	3830462450

We have 26 States in Brazil, now we just have to find out how many different products we have to think about the best representation.

62

105237	Soja, mesmo triturada, exceto para semeadura
105174	Carnes desossadas de bovino, congeladas
105184	Milho em grão, exceto para semeadura
65022	Carnes desossadas de bovino, frescas ou refrig...
105141	Óleos brutos de petróleo
	...
30096	Outros aparelhos para filtrar ou depurar líquidos
65158	Outras madeiras compensadas, constituídas excl...
65095	Castanha-do-pará, fresca ou seca, com casca
65161	Outras miudezas comestíveis de bovino, congeladas
39	Outras madeiras de não coníferas perfilada (co...

Name: Produto, Length: 62, dtype: object

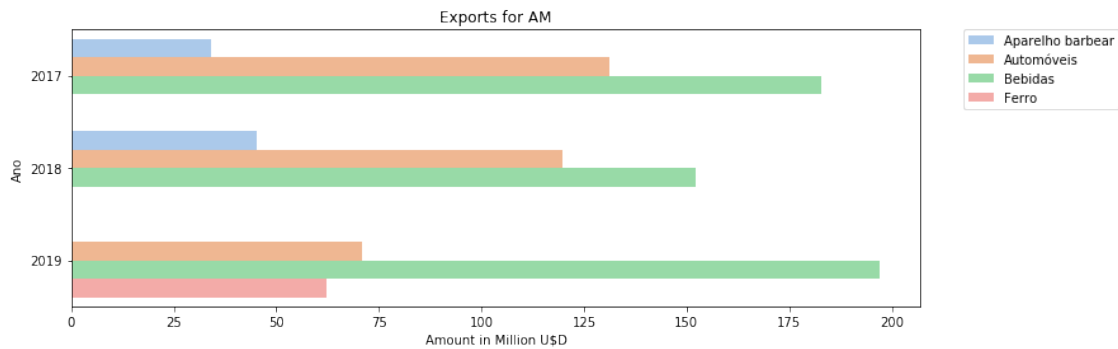
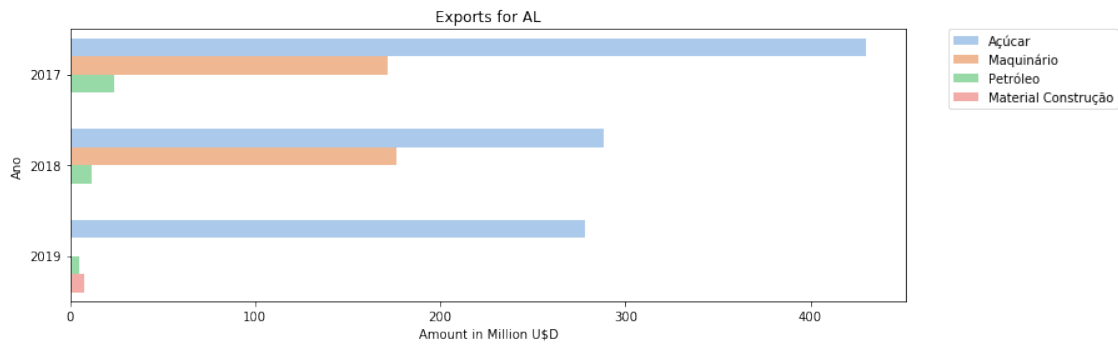
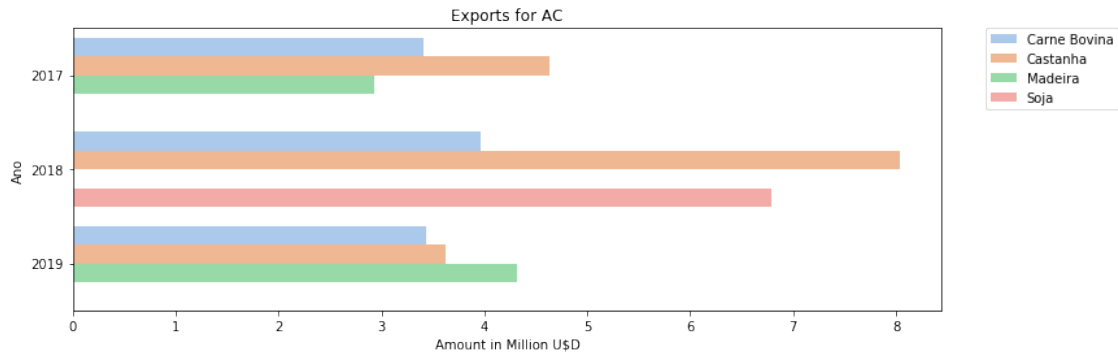
The names are hard to understand and there are many similar products in different categories. For this reason, we will rename them and merge when necessary.

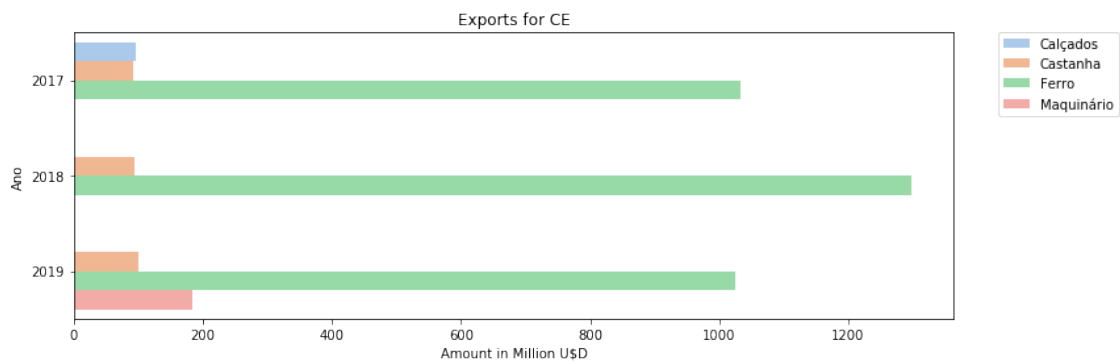
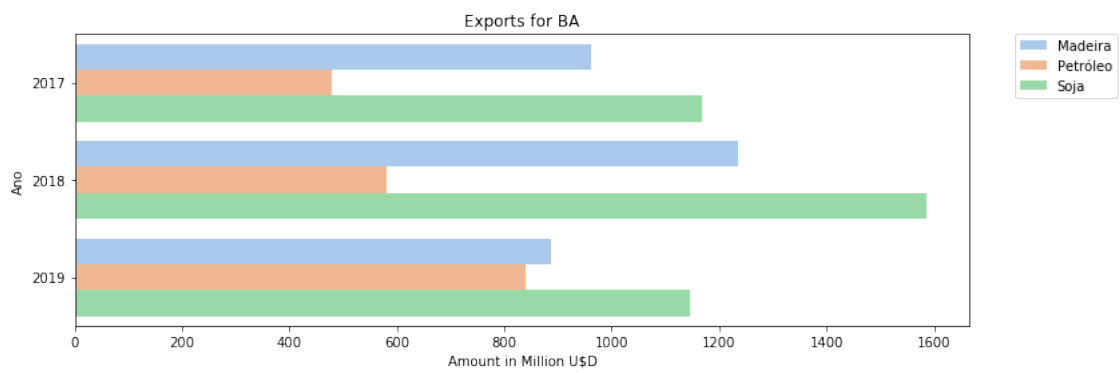
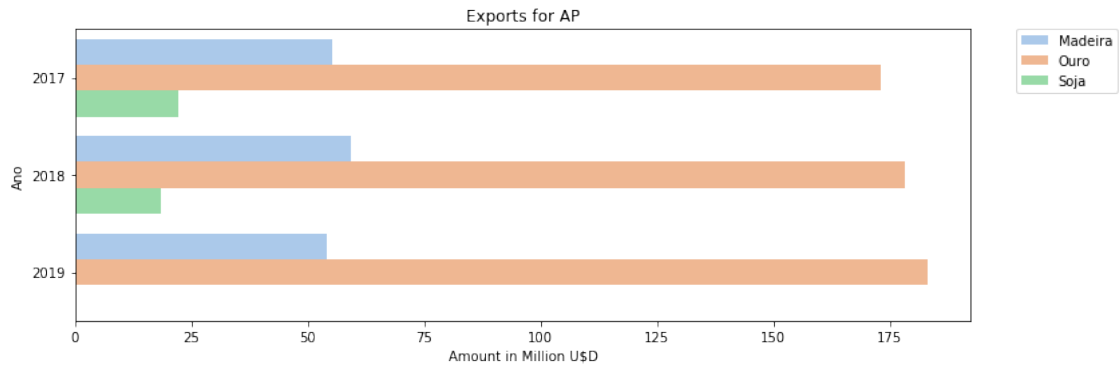
	Ano	Estado	Produto	Amount (USD)
0	2019	TO	Soja	772631871
1	2019	TO	Carne Bovina	155101002
2	2019	TO	Milho	78543244
3	2018	TO	Soja	995302750
4	2018	TO	Carne Bovina	88538892
5	2018	TO	Carne Bovina	25932145
6	2017	TO	Soja	755967278
7	2017	TO	Carne Bovina	78877967
8	2017	TO	Milho	53312114
9	2019	SP	Petróleo	3830462450

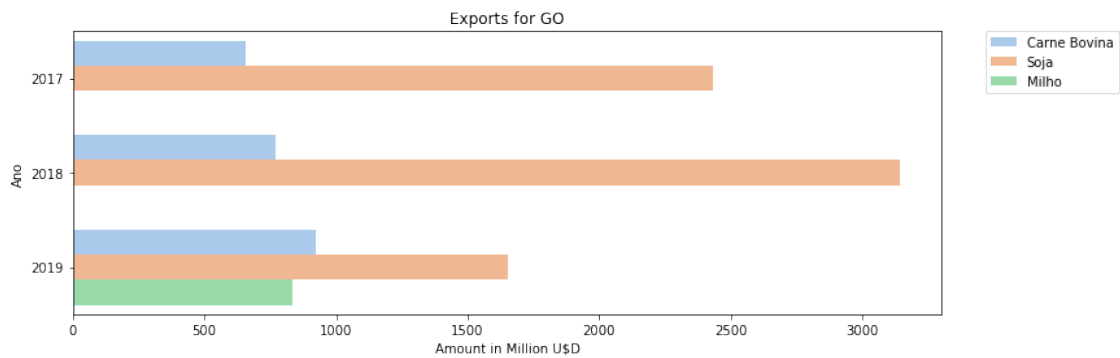
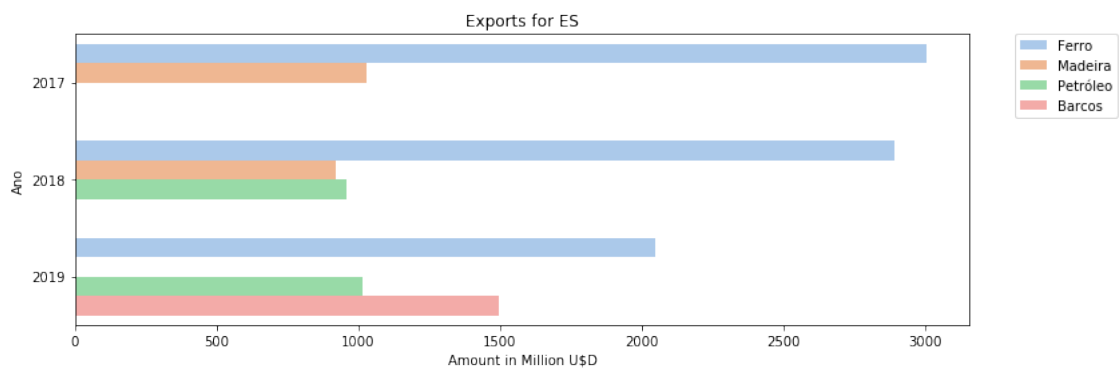
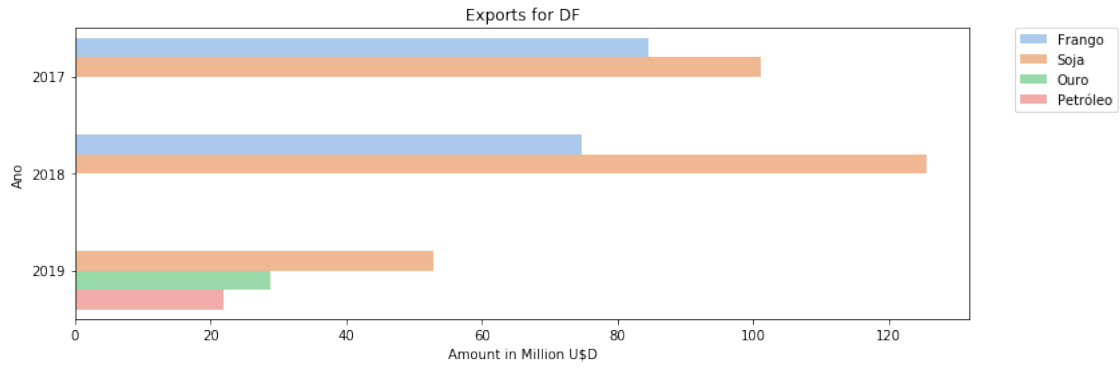
Now we have another problem: due to the renaming, two different products in our original table might have become one single product. For an example: soy beans and soy flour are now just Soy. For this reason, we have to make another groupby and sum.

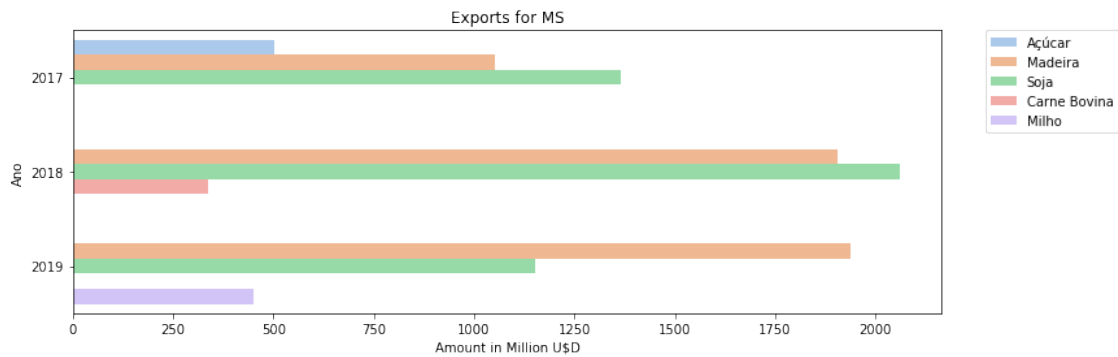
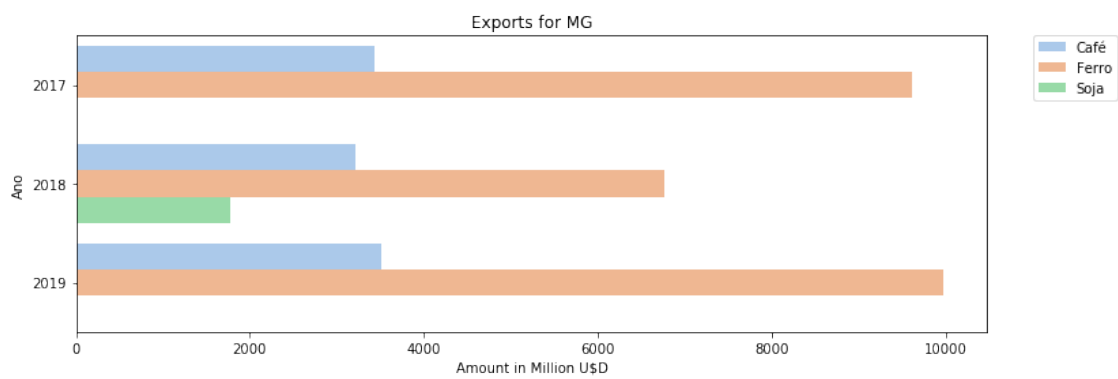
	Ano	Estado	Produto	Amount (USD)
0	2017	AC	Carne Bovina	3.413598
1	2017	AC	Castanha	4.631682
2	2017	AC	Madeira	2.932218
3	2017	AL	Açúcar	430.227643
4	2017	AL	Maquinário	171.657240

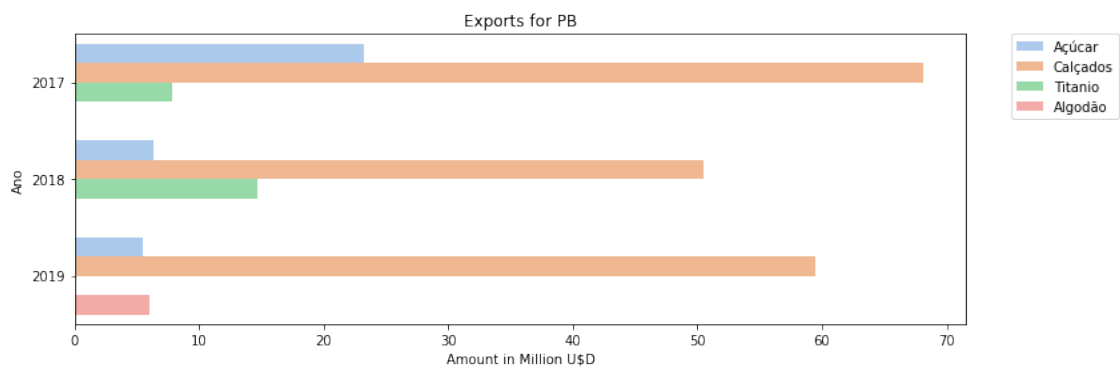
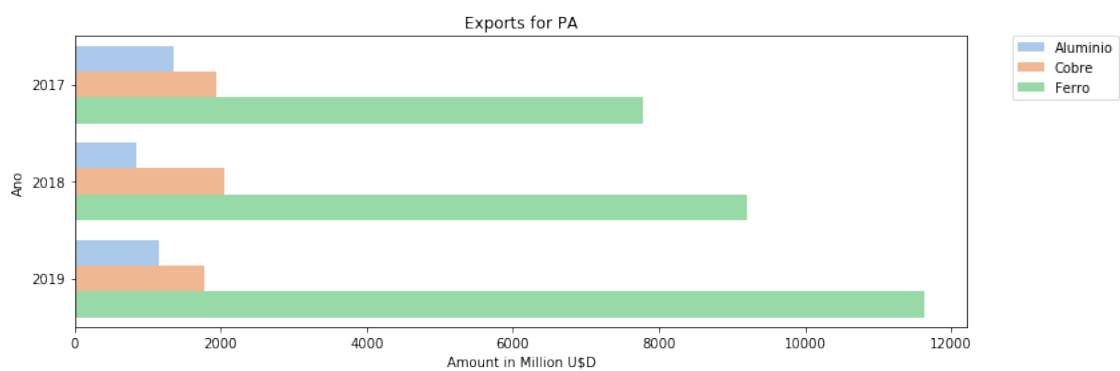
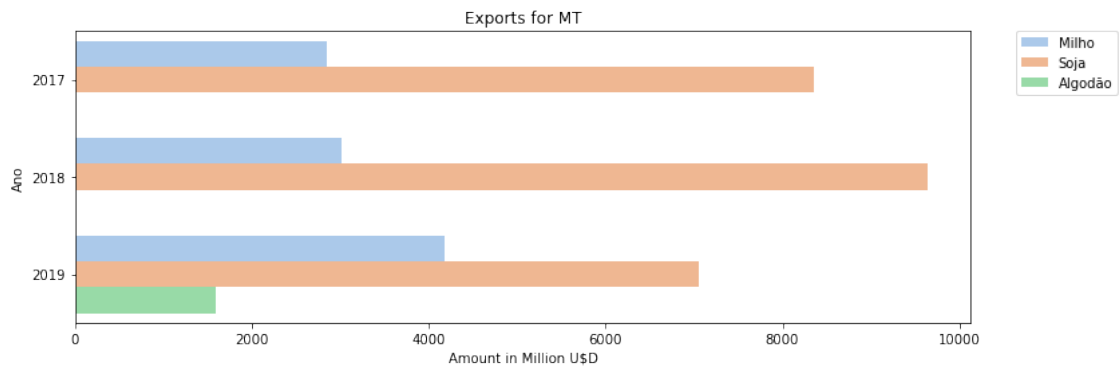
5	2017	AL	Petróleo	23.815680
6	2017	AM	Aparelho barbear	34.155710
7	2017	AM	Automóveis	131.141383
8	2017	AM	Bebidas	182.978114
9	2017	AP	Madeira	55.283268

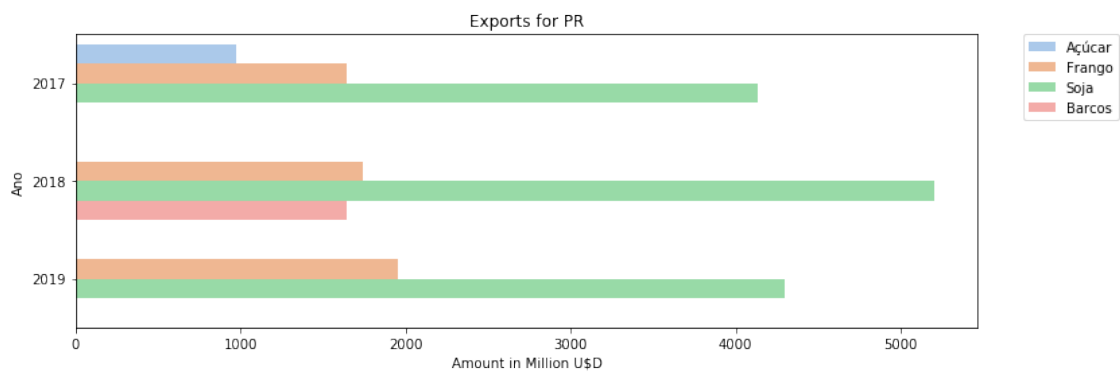
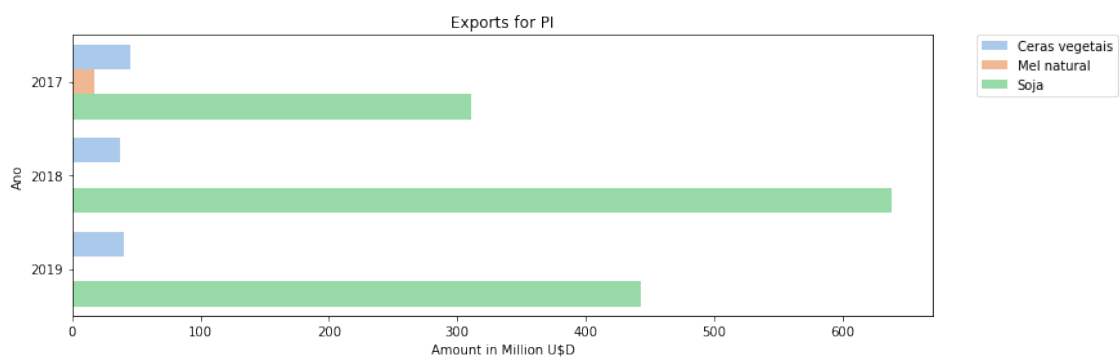
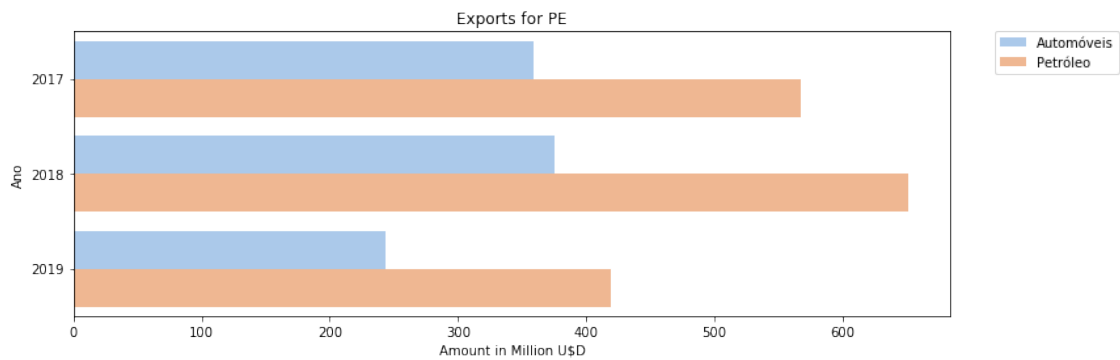


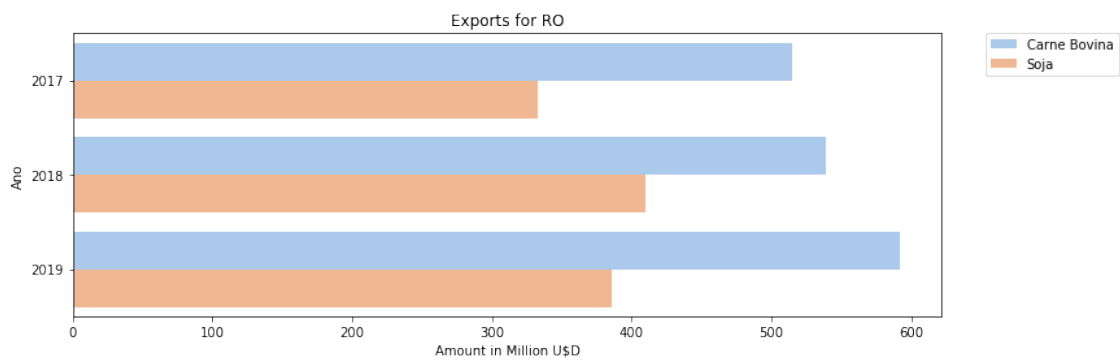
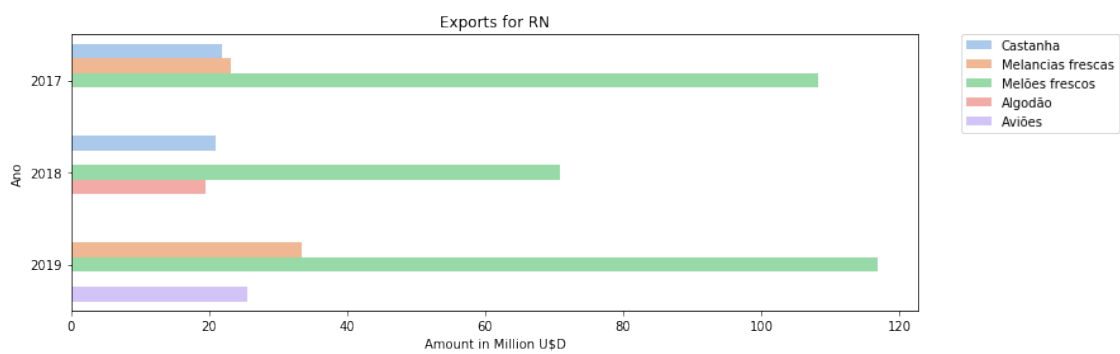
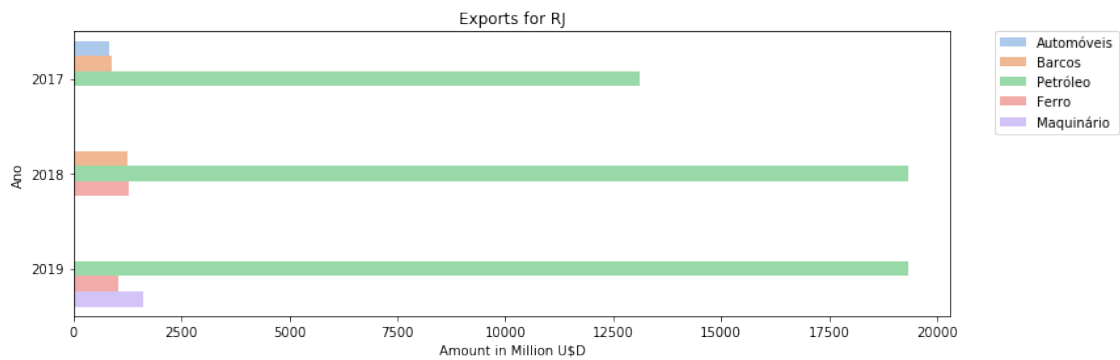


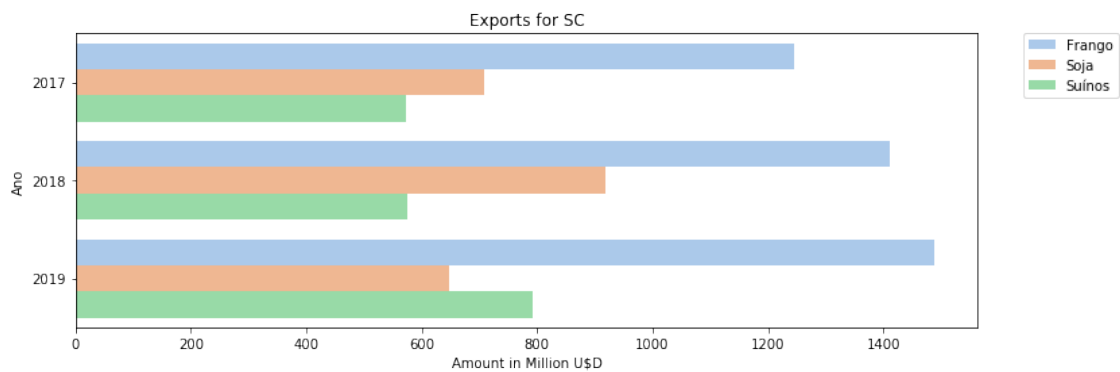
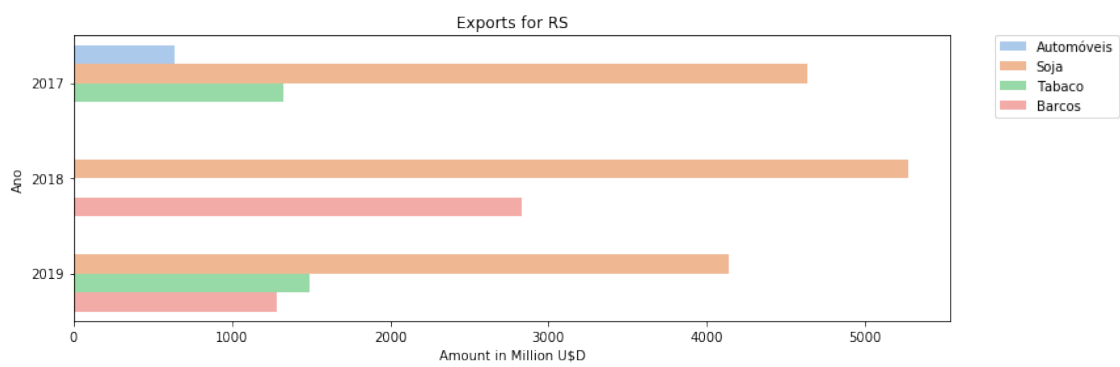
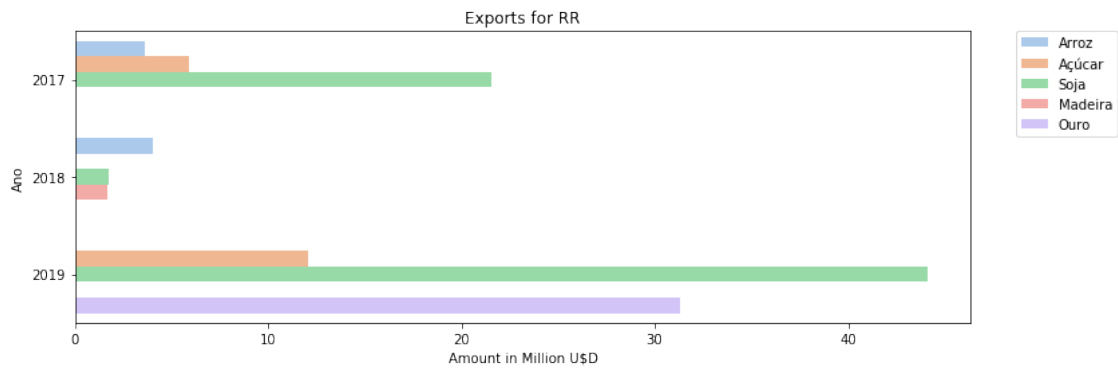


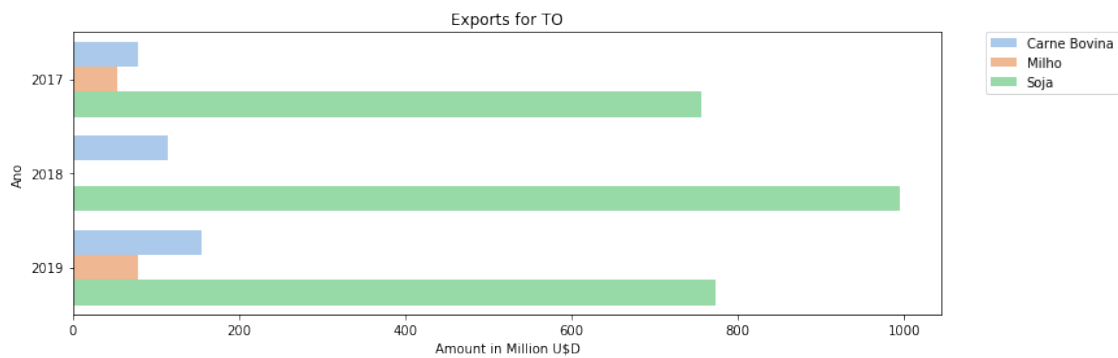
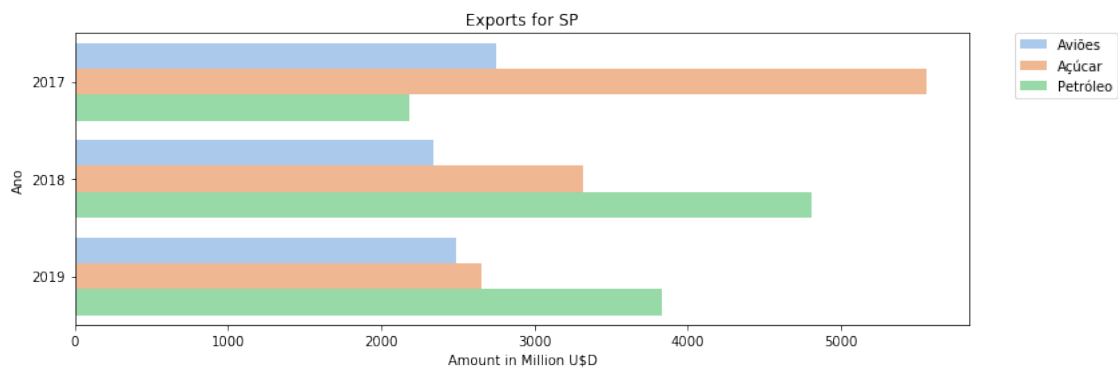
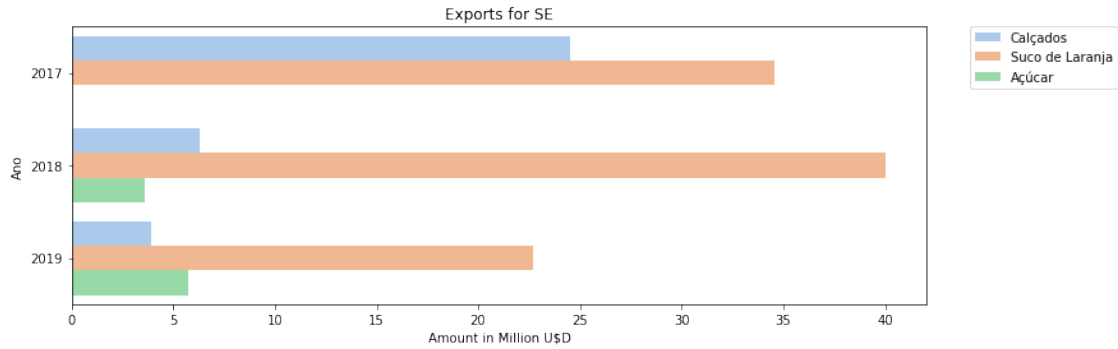








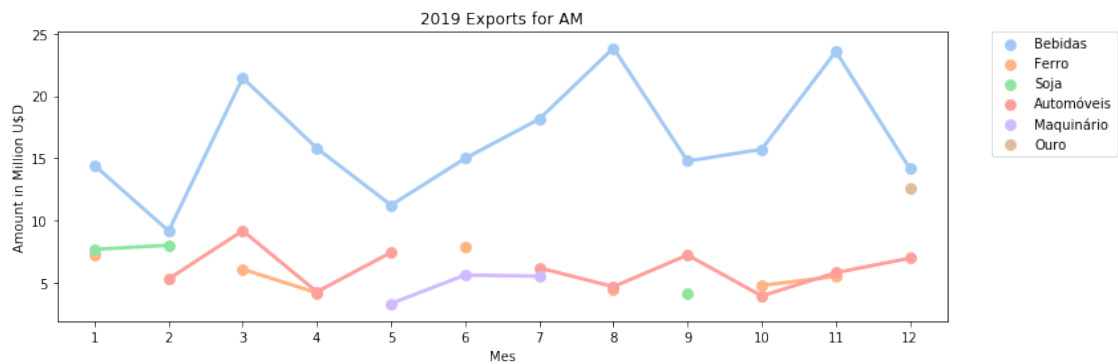
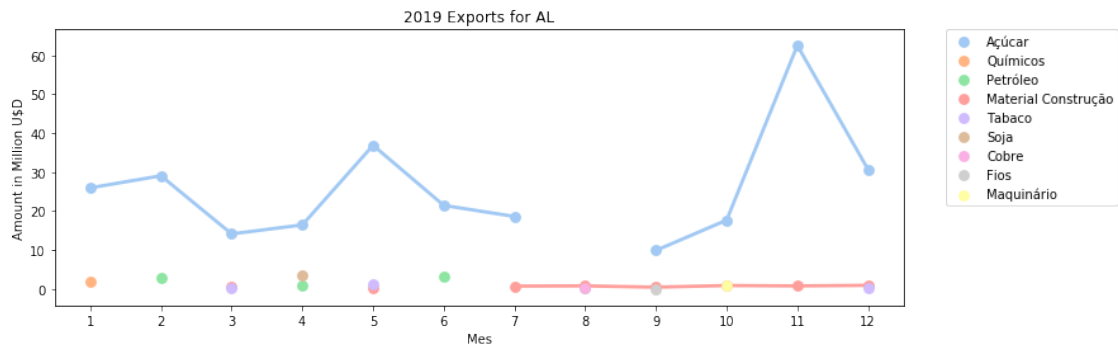
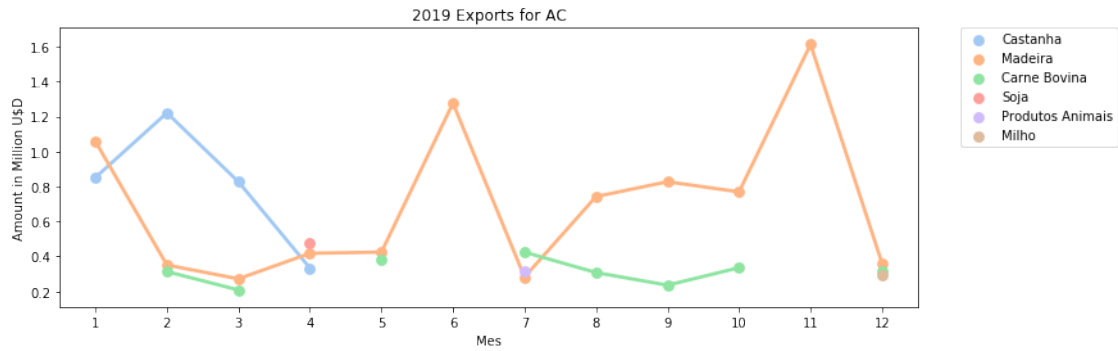


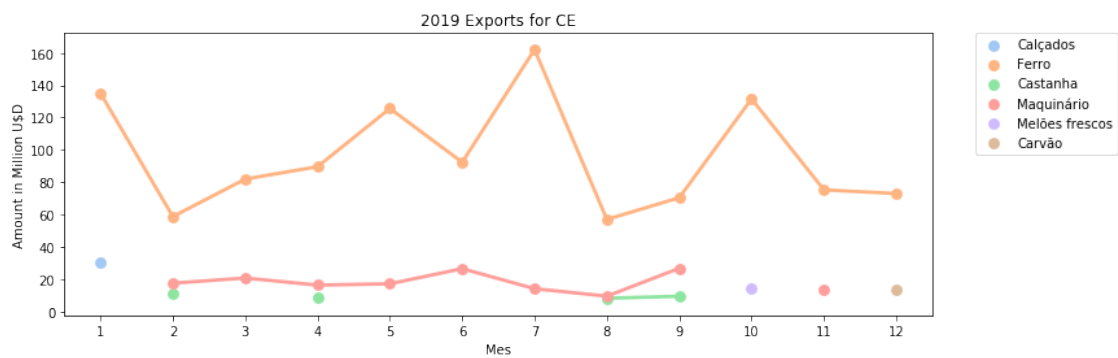
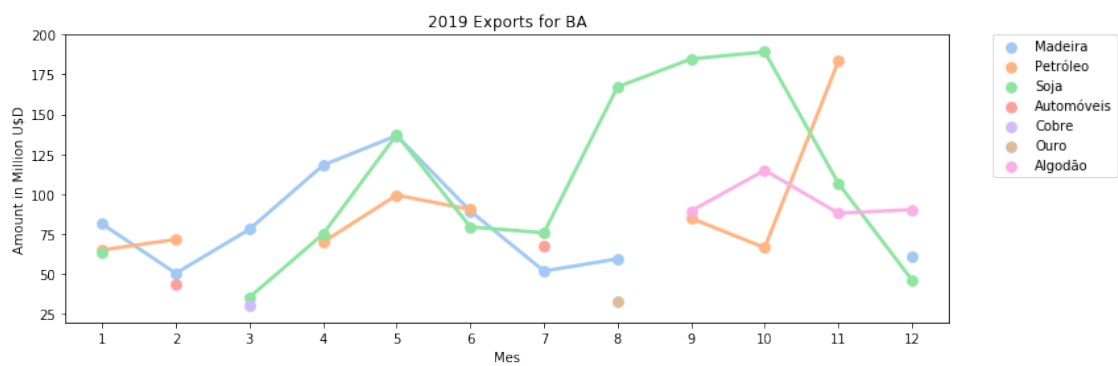
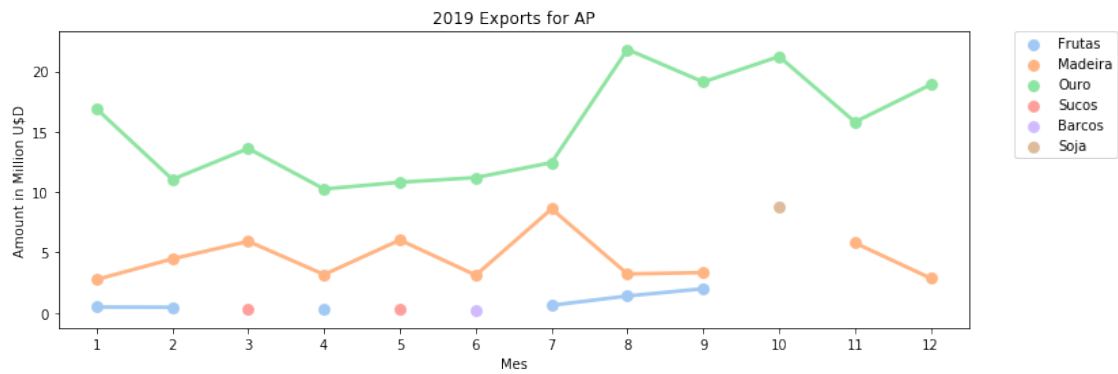


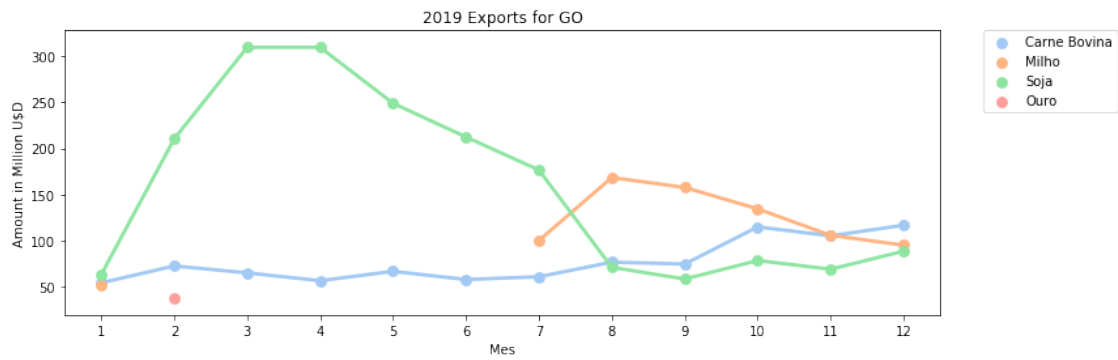
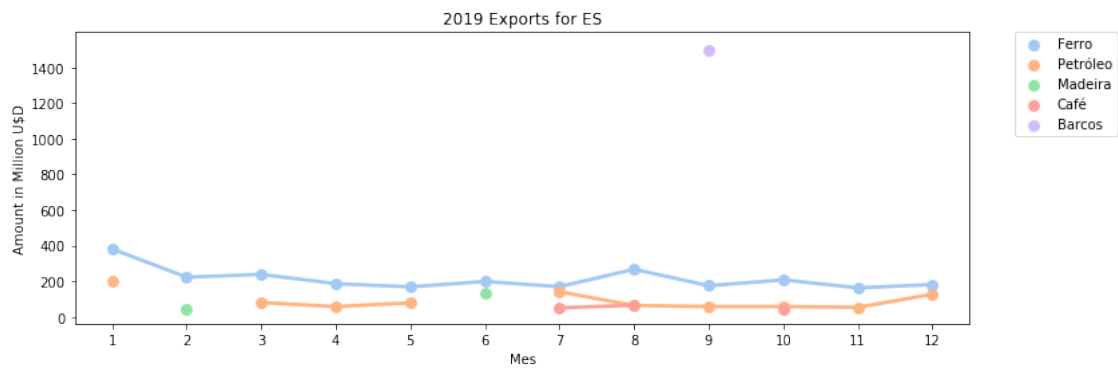
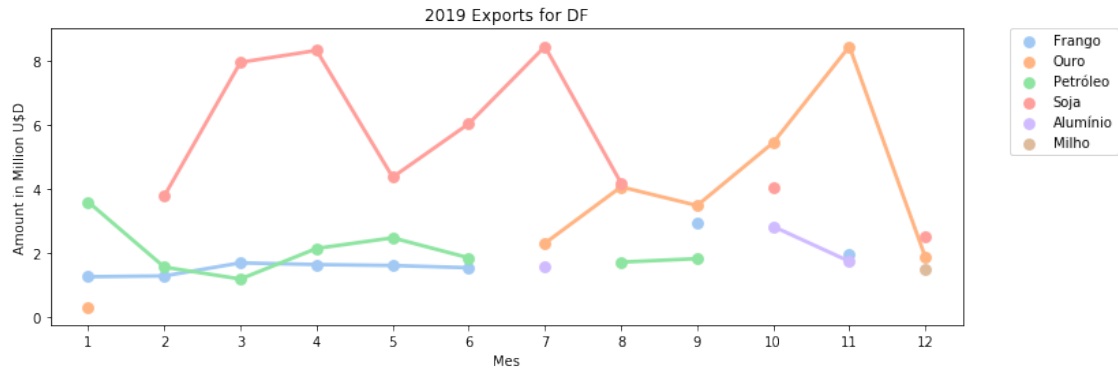
We notice some States had only 2 exports, and this means that after the name change to make visualization easier, two products merged into one.

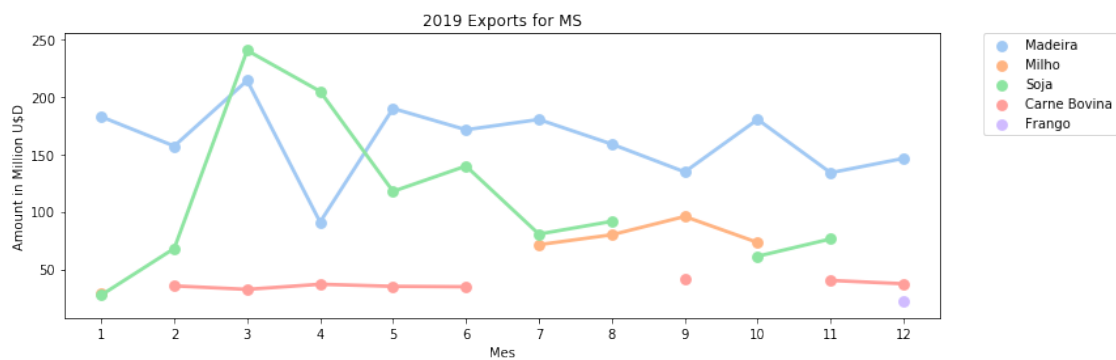
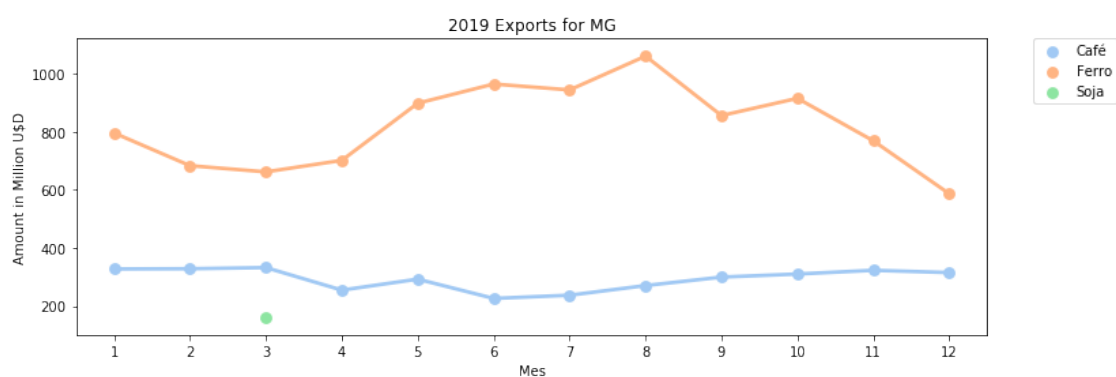
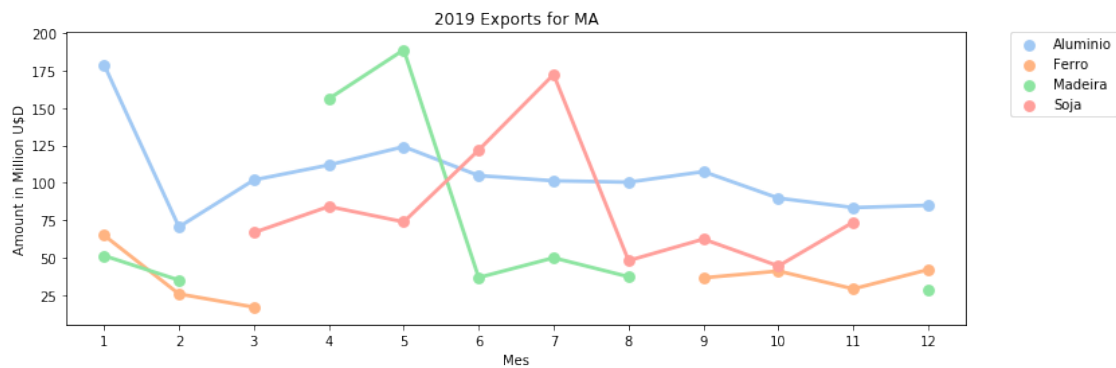
	Ano	Mes	Estado	Produto	Amount(USD)
0	2019	1	AC	Castanha	0.851810
1	2019	1	AC	Madeira	1.058688
2	2019	1	AL	Açúcar	25.985332
3	2019	1	AL	Químicos	1.806044

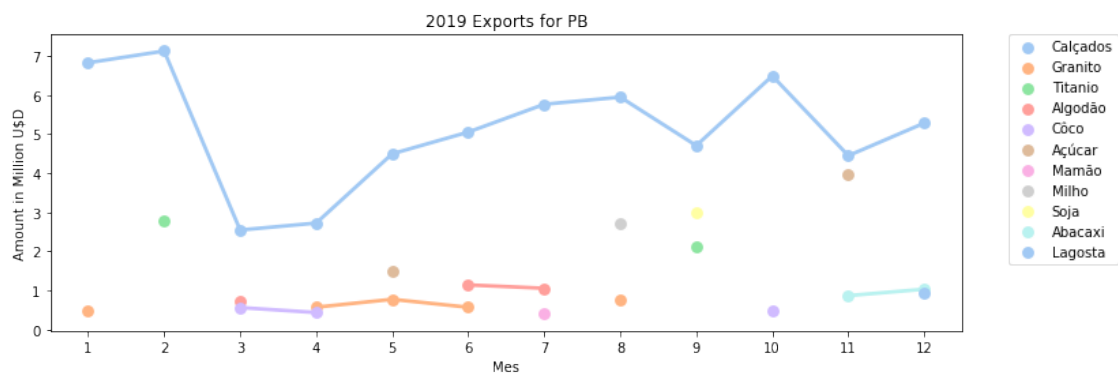
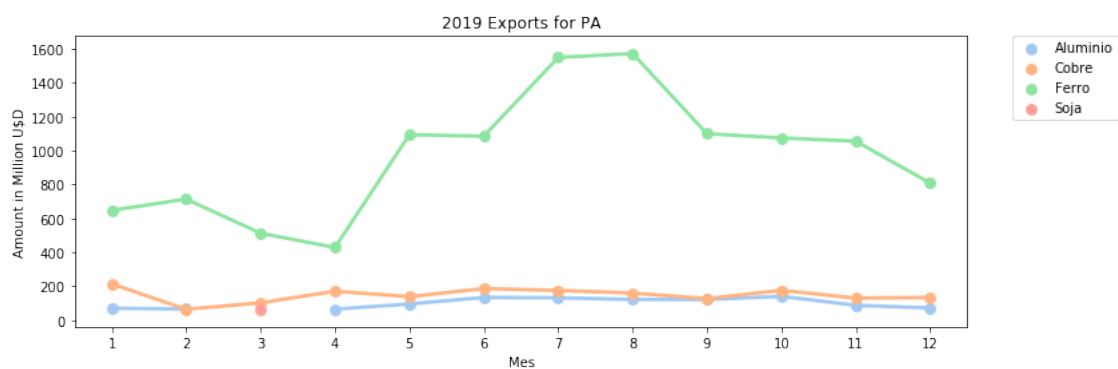
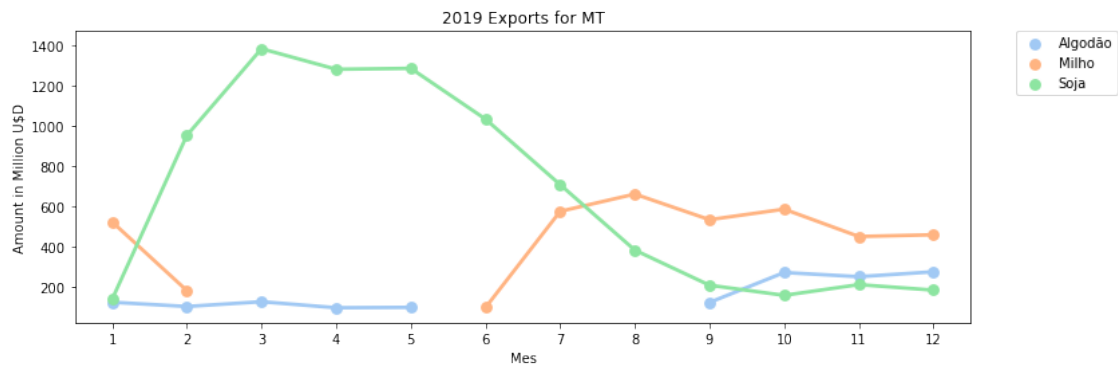
4	2019	1	AM	Bebidas	14.425506
5	2019	1	AM	Ferro	7.253336
6	2019	1	AM	Soja	7.677951
7	2019	1	AP	Frutas	0.484048
8	2019	1	AP	Madeira	2.761465
9	2019	1	AP	Ouro	16.900377

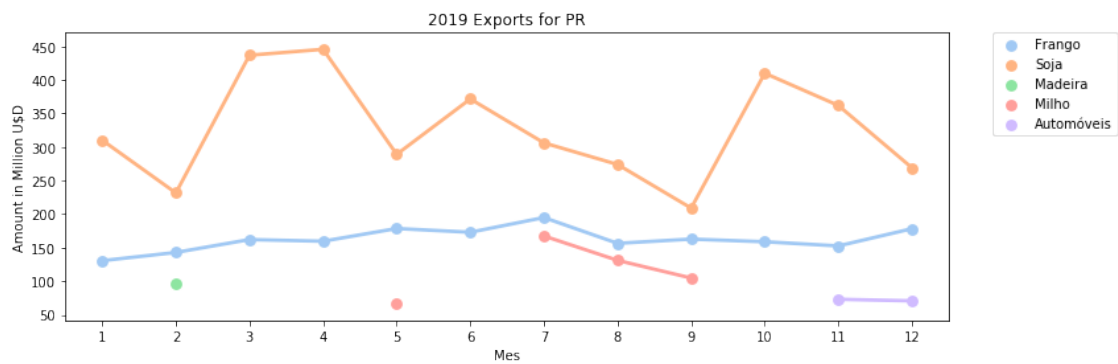
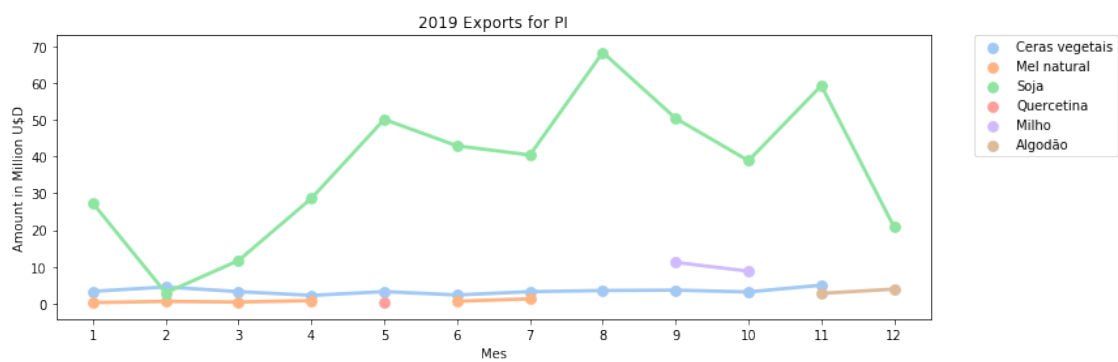
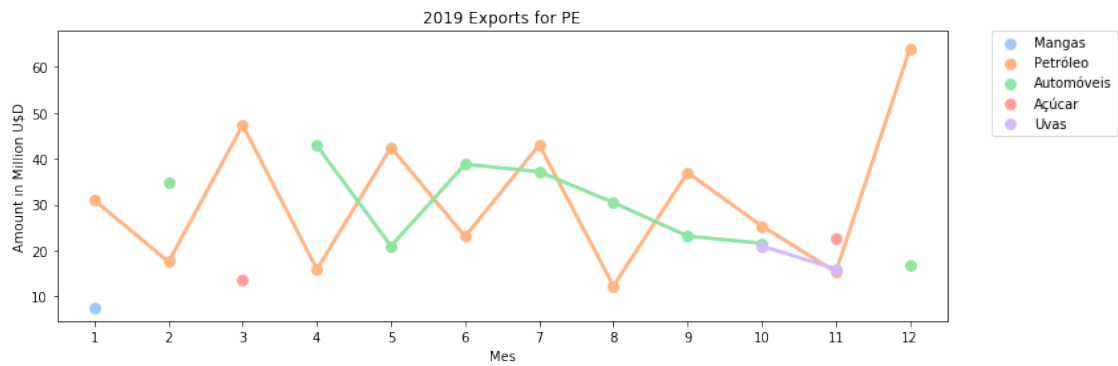


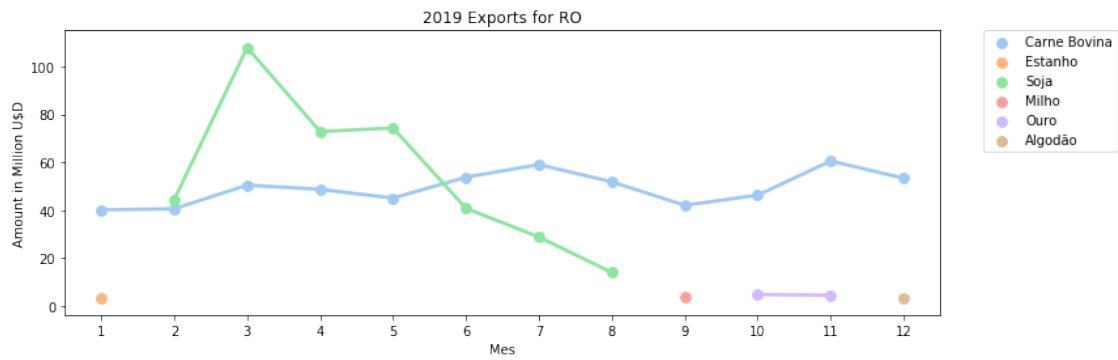
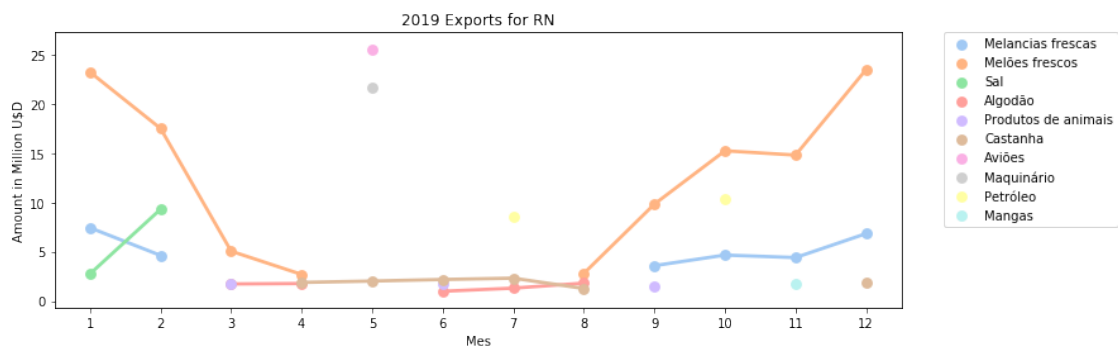
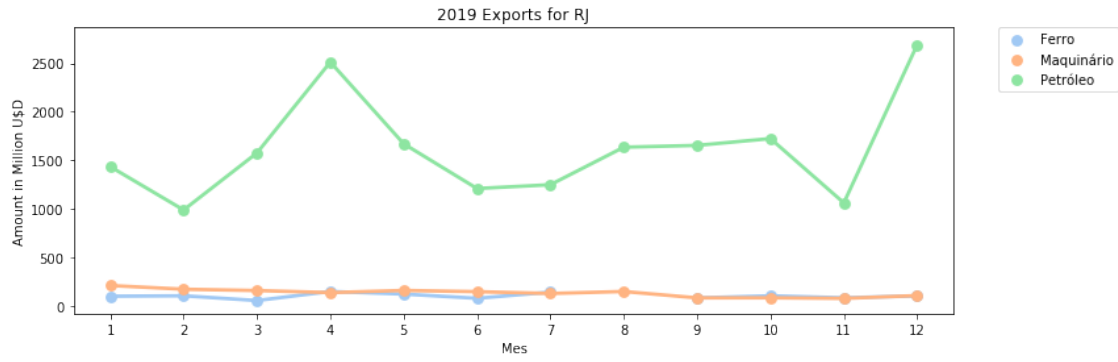


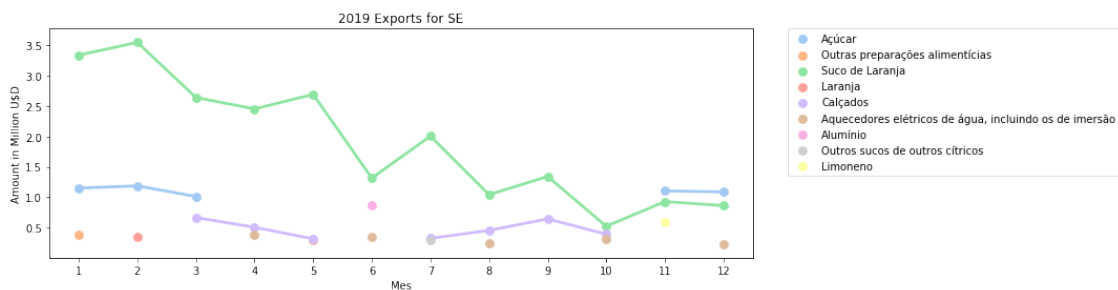
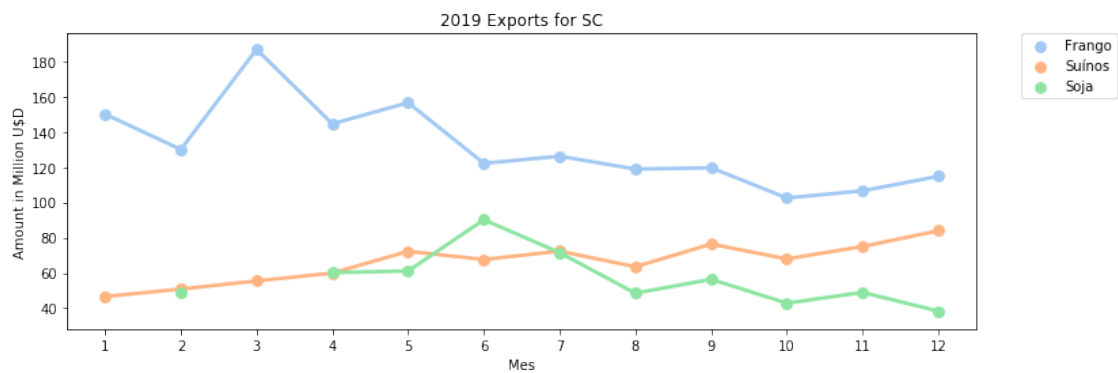
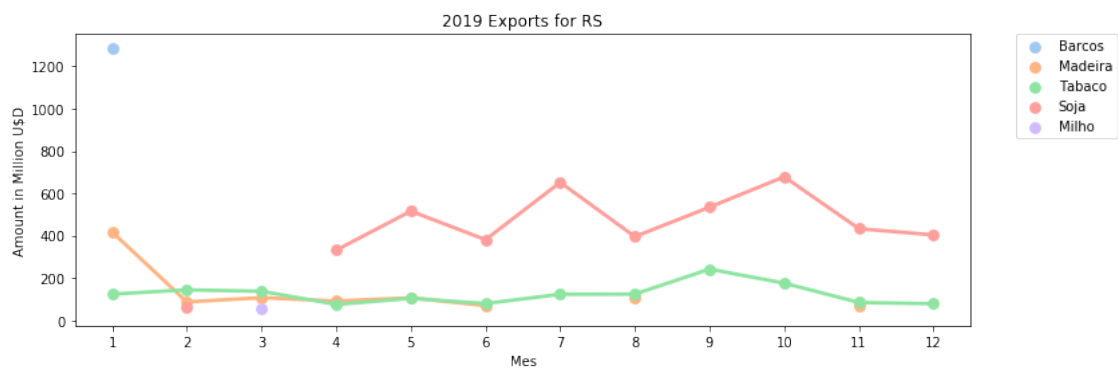
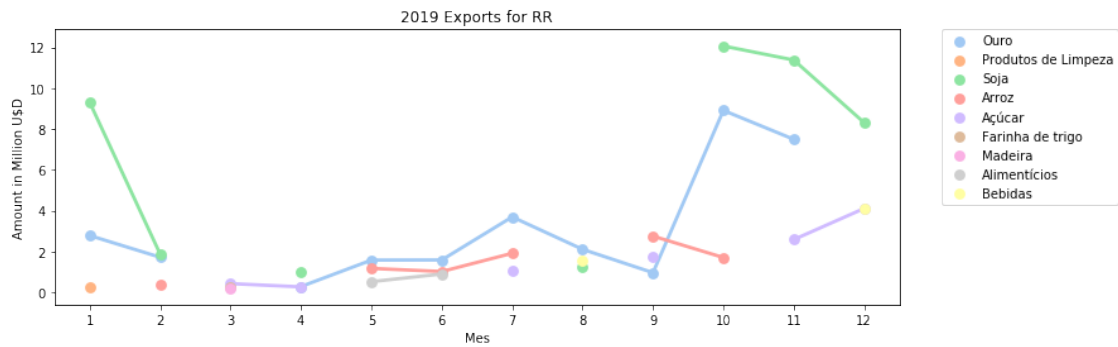


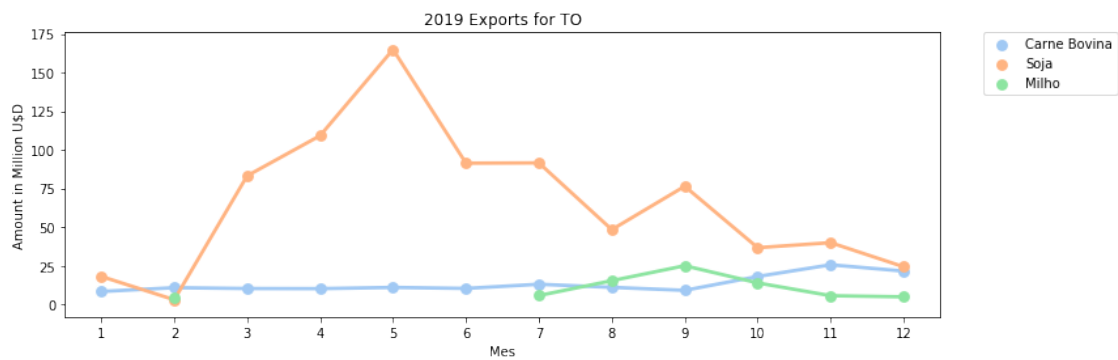
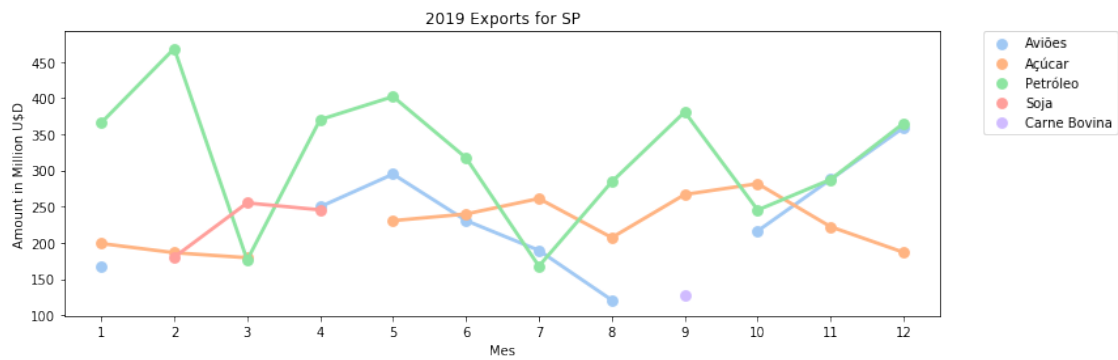












56

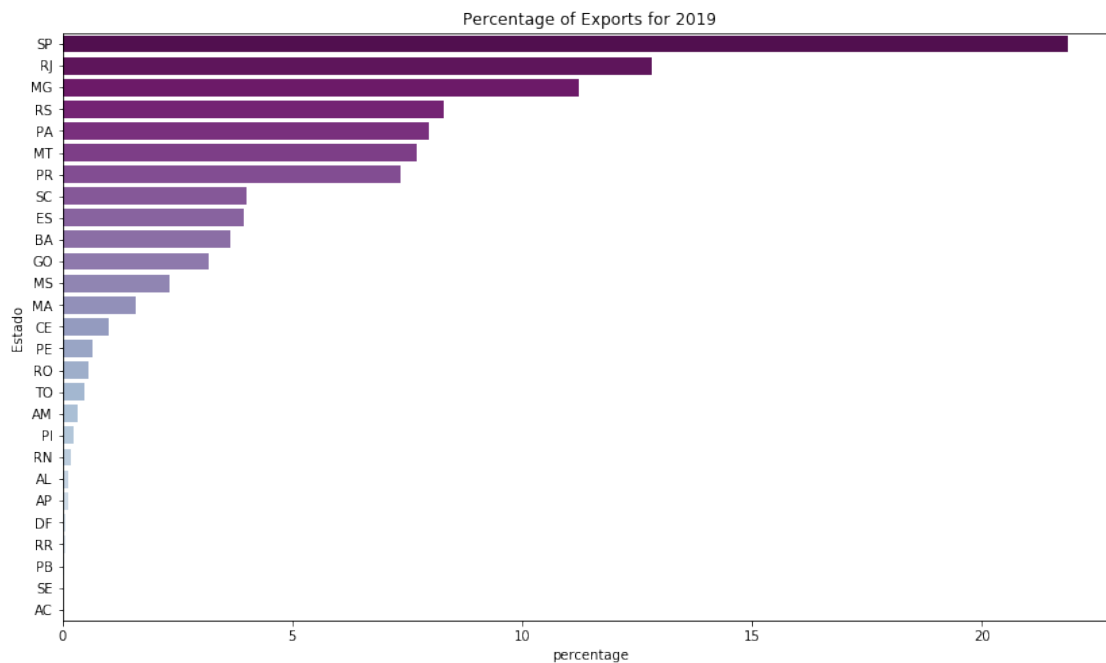
	Ano	Estado	Amount(USD)
0	2019	AC	32853291
1	2019	AL	319088835
2	2019	AM	731091968
3	2019	AP	261368366
4	2019	BA	8168158116
5	2019	CE	2275188077
6	2019	DF	160701336
7	2019	ES	8800321849
8	2019	GO	7133398211
9	2019	MA	3543622779

223446289725

	Ano	Estado	Amount(USD)	percentage
25	2019	SP	48852560073	21.863223

18	2019	RJ	28634458633	12.814918
10	2019	MG	25138578745	11.250390
22	2019	RS	18545065307	8.299563
13	2019	PA	17841239155	7.984576
12	2019	MT	17206103910	7.700331
17	2019	PR	16454197120	7.363827
23	2019	SC	8951838846	4.006260
7	2019	ES	8800321849	3.938451
4	2019	BA	8168158116	3.655535

Text(0.5, 1.0, 'Percentage of Exports for 2019')



1.4 Data Wrangling for Imports

Now we load and handle the imports file.

	Ano	Mes	Estado	Pais	Via	\
0	2017	1	SP	Japão	AEREA	
1	2017	1	AM	Japão	AEREA	
2	2017	9	RS	Japão	AEREA	
3	2017	2	SP	Japão	AEREA	
4	2017	7	BA	Japão	AEREA	
5	2017	10	PR	Japão	AEREA	
6	2017	4	SP	Japão	AEREA	
7	2017	9	SP	Japão	AEREA	

8	2017	2	AM	Japão	AEREA
9	2017	8	RS	Japão	AEREA

	Produto	Kg	Amount (USD)
0	Condensador fixo elétrico, de tântalo, próprio...	99	60920
1	Condensador fixo elétrico, de tântalo, próprio...	5	6876
2	Condensador fixo elétrico, de tântalo, próprio...	0	322
3	Condensador fixo elétrico, de tântalo, próprio...	26	37233
4	Condensador fixo elétrico, de tântalo, próprio...	0	22
5	Condensador fixo elétrico, de tântalo, próprio...	0	111
6	Condensador fixo elétrico, de tântalo, próprio...	0	169
7	Condensador fixo elétrico, de tântalo, próprio...	135	120802
8	Condensador fixo elétrico, de tântalo, próprio...	3	2170
9	Condensador fixo elétrico, de tântalo, próprio...	1	302

Ano	0
Mes	0
Estado	0
Pais	0
Via	0
Produto	0
Kg	0
Amount (USD)	0

dtype: int64

0	SP
1	AM
2	RS
4	BA
5	PR
10	MG
12	GO
17	SC
45	PE
168	RJ
177	ES
374	MS
511	DF
520	AL
674	RO
739	CE
759	MA
795	PA
2842	RN
3344	PB
3407	SE
3993	MT

```

4349    PI
4687    ZN
5474    AP
11344   TO
11804   RR
28100   AC
Name: Estado, dtype: object

```

```

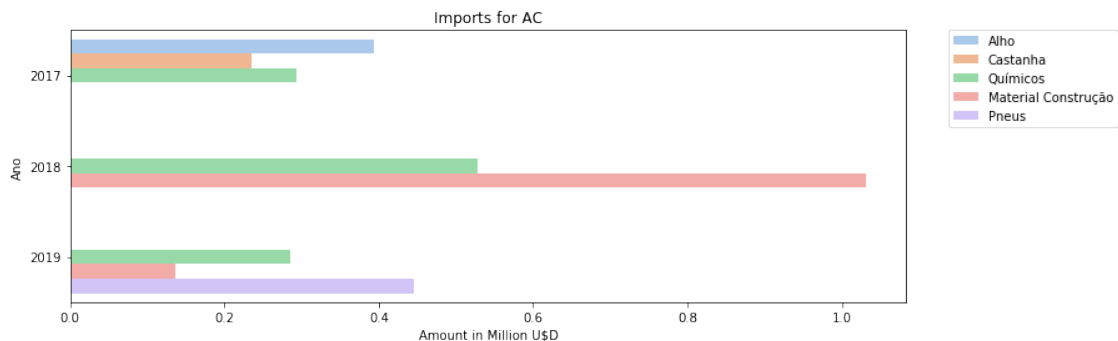
0          Japão
242        França
244        Áustria
248        China
612        Alemanha
...
3081654    Togo
3092881    Wake, Ilha
3231069    São Pedro e Miquelon
3650893    Heard e ilhas mcdonald, Ilha
5544590    Antártica
Name: Pais, Length: 250, dtype: object

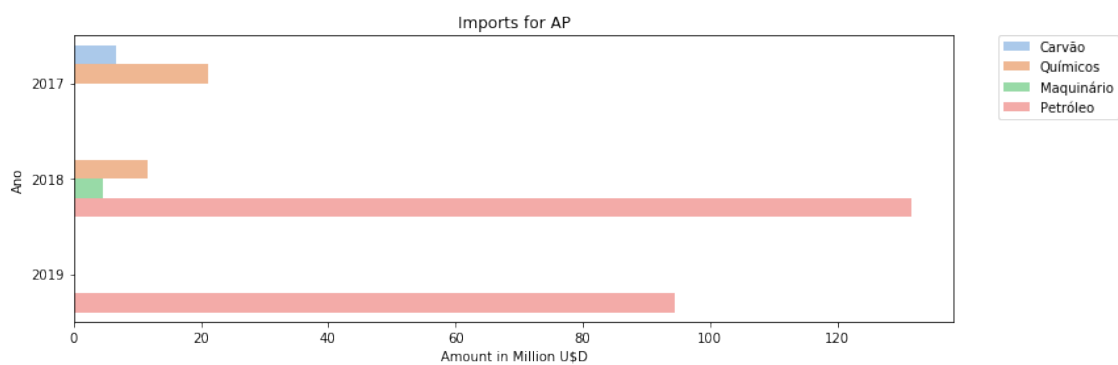
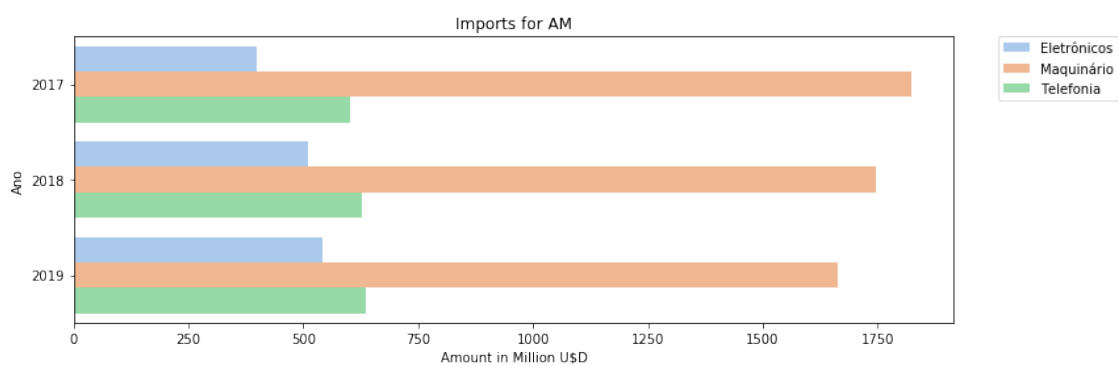
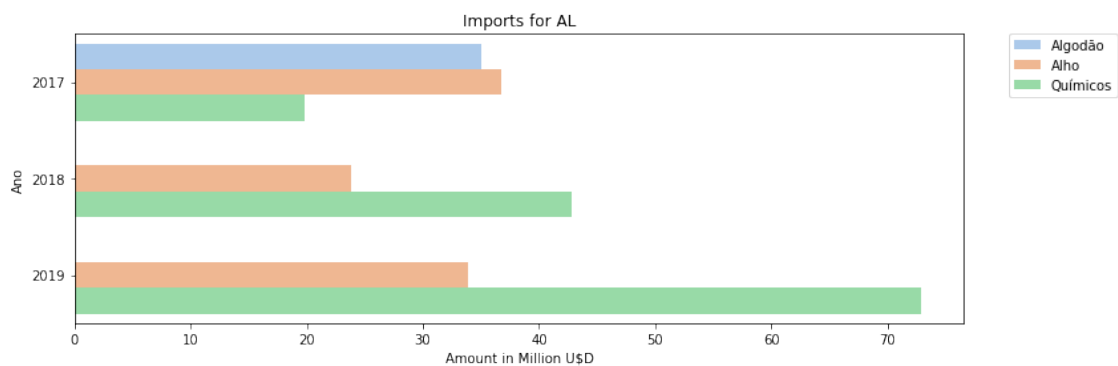
```

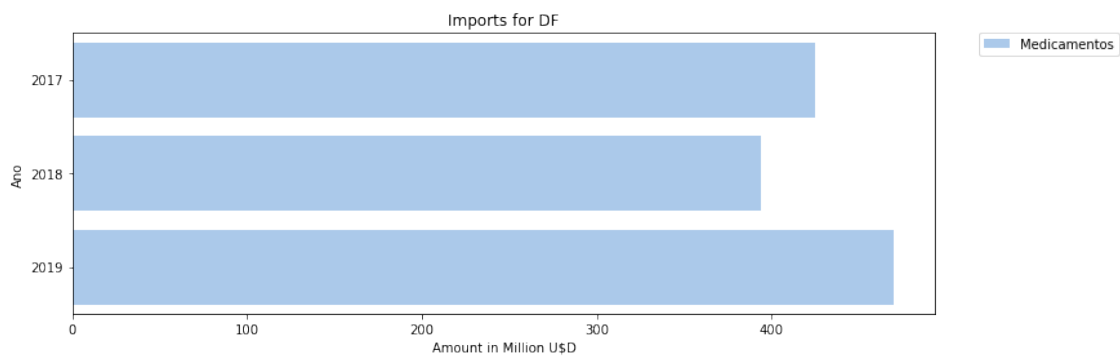
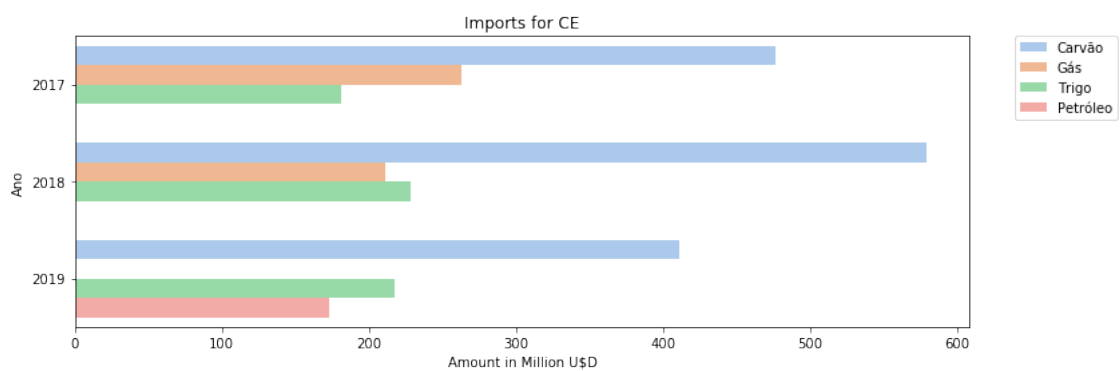
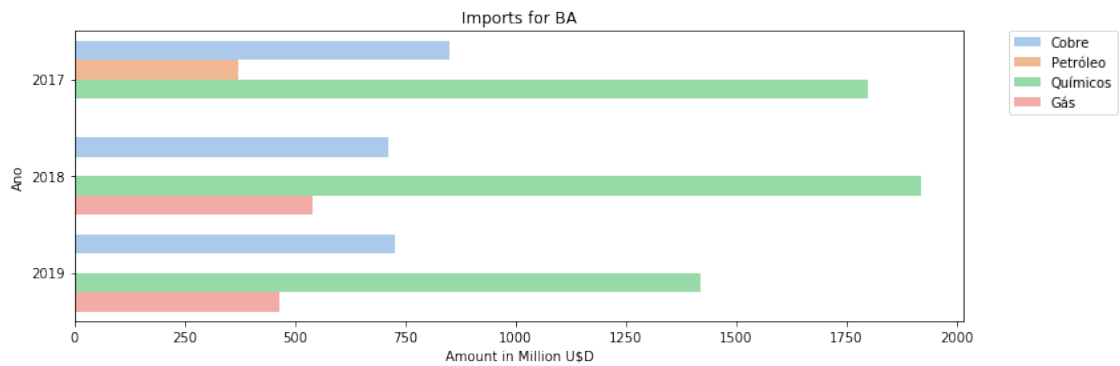
1.5 Imports Analysis

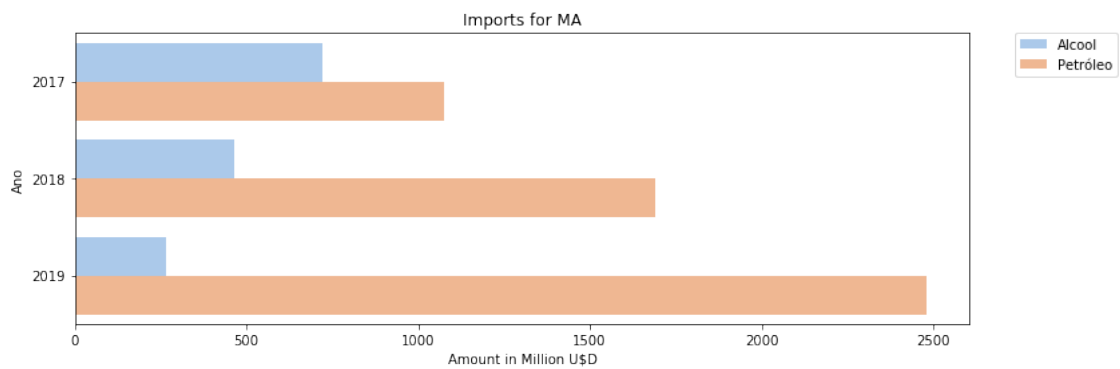
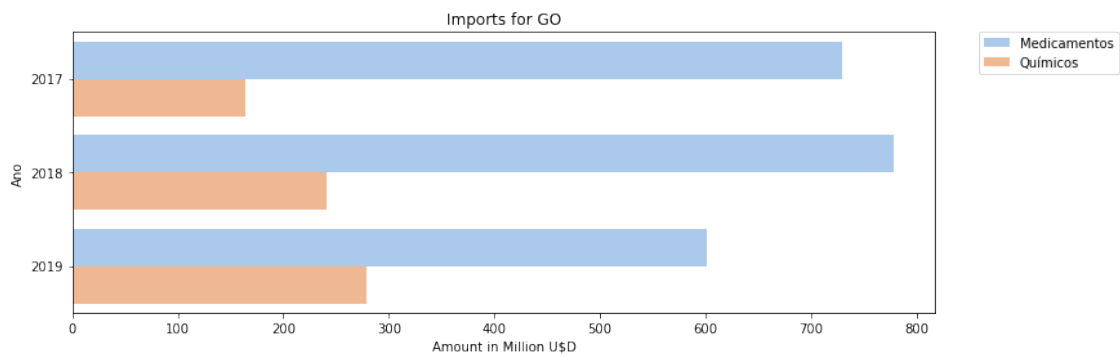
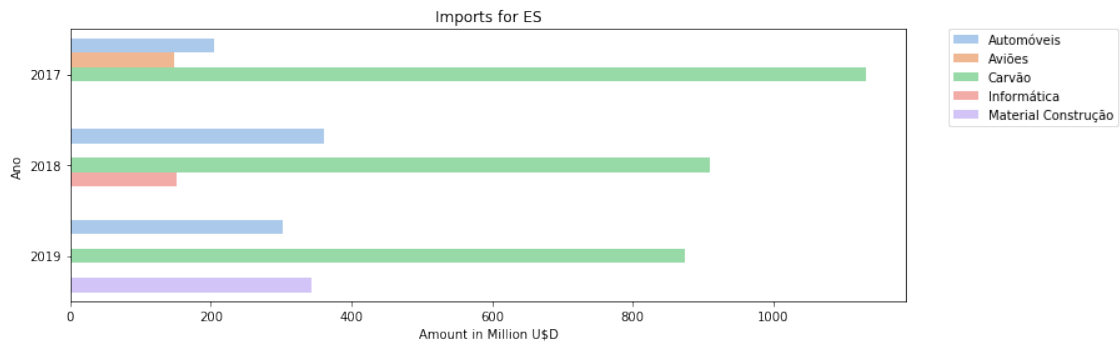
Now it is time to analyse our imports.

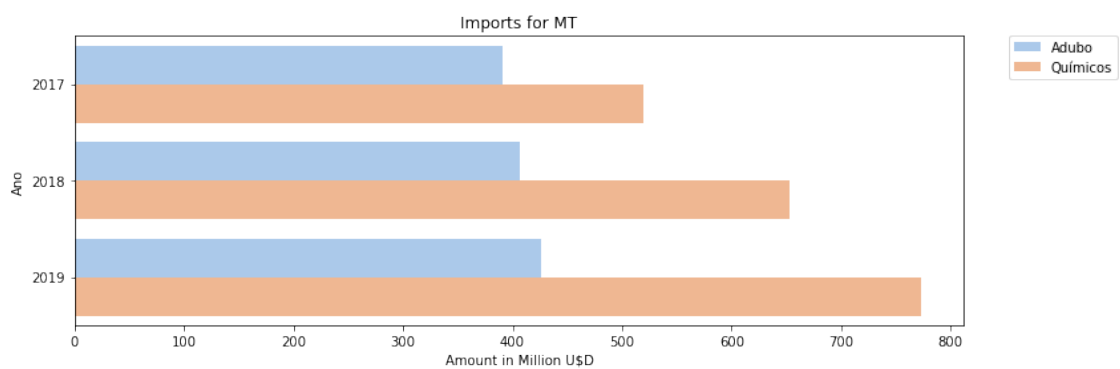
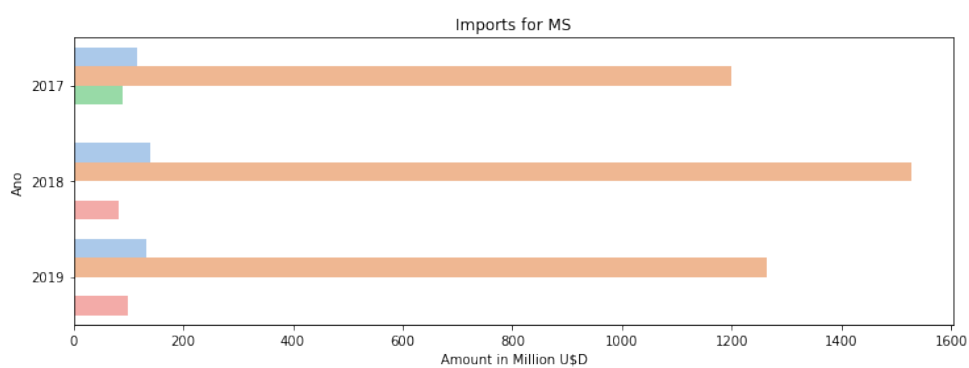
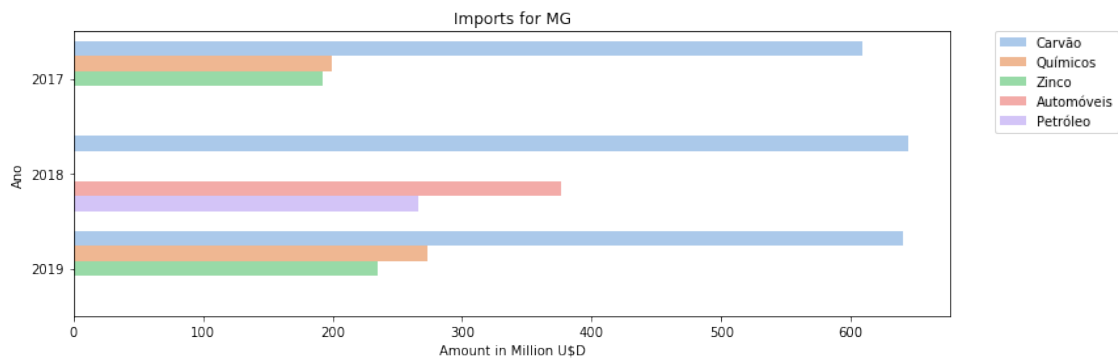
	Ano	Estado	Produto	Amount(USD)
0	2017	AC	Alho	0.393150
1	2017	AC	Castanha	0.235580
2	2017	AC	Químicos	0.293144
3	2017	AL	Algodão	35.044473
4	2017	AL	Alho	36.706566
5	2017	AL	Químicos	19.865314
6	2017	AM	Eletrônicos	397.652474
7	2017	AM	Maquinário	1824.483948
8	2017	AM	Telefonia	600.288454
9	2017	AP	Carvão	6.622802

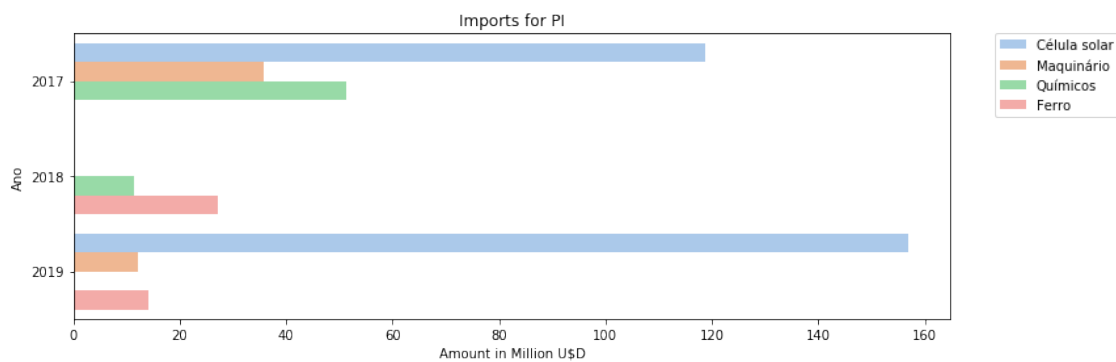
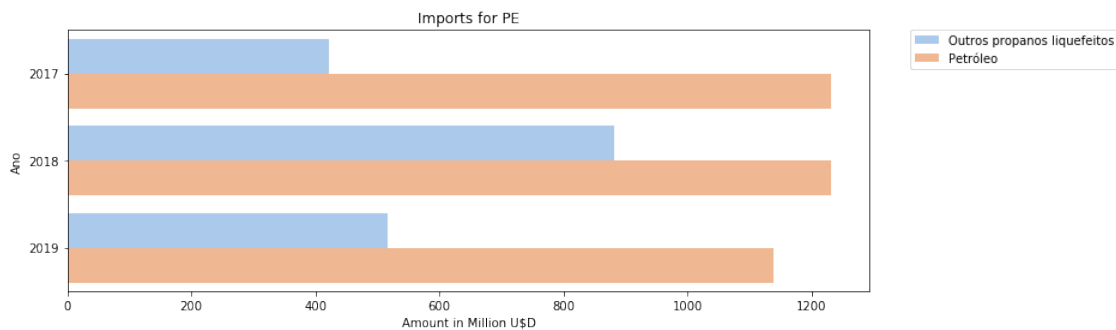
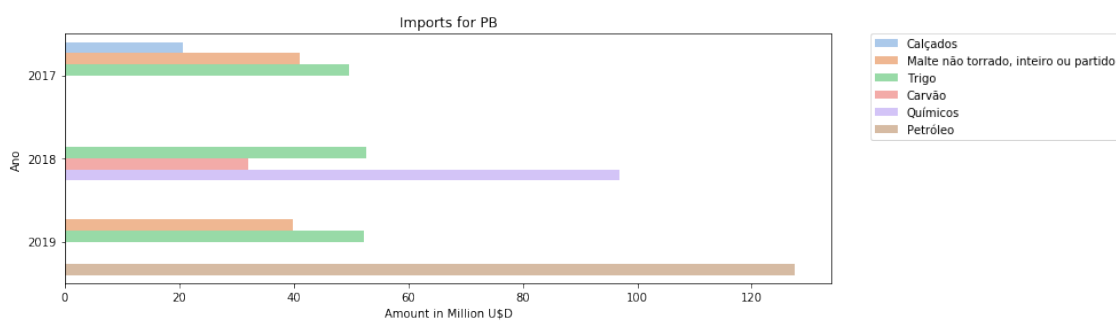
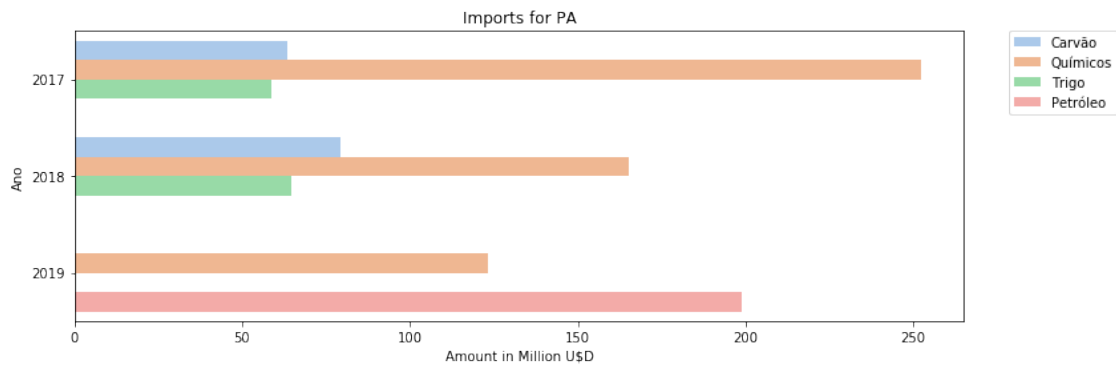


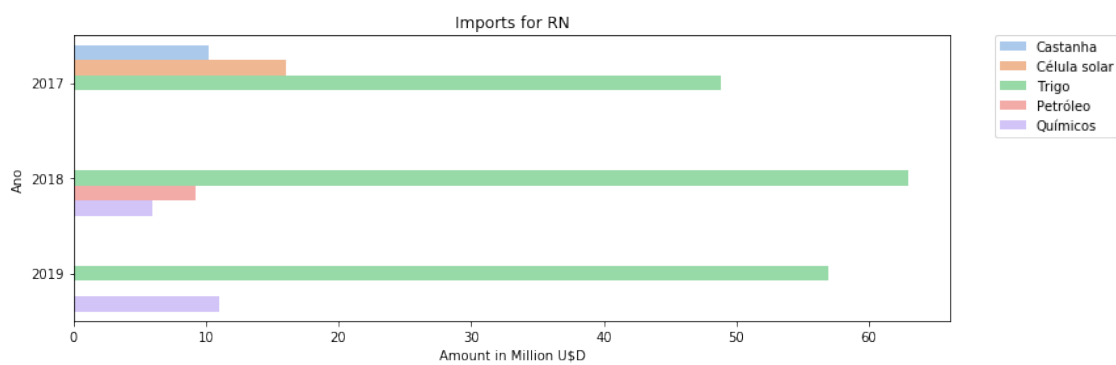
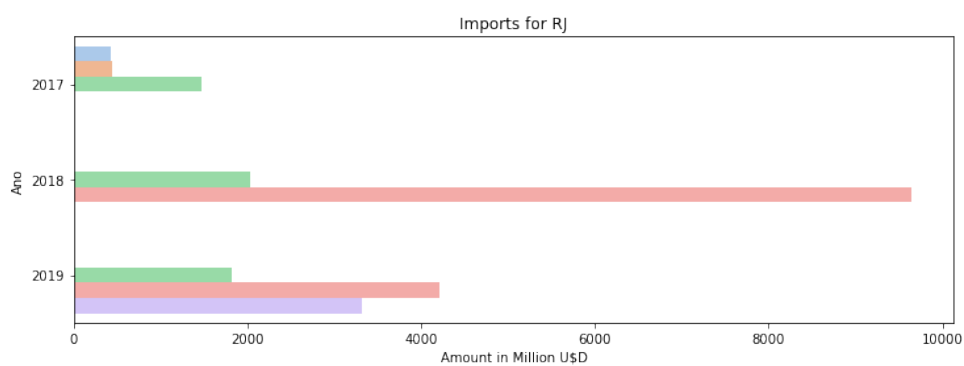
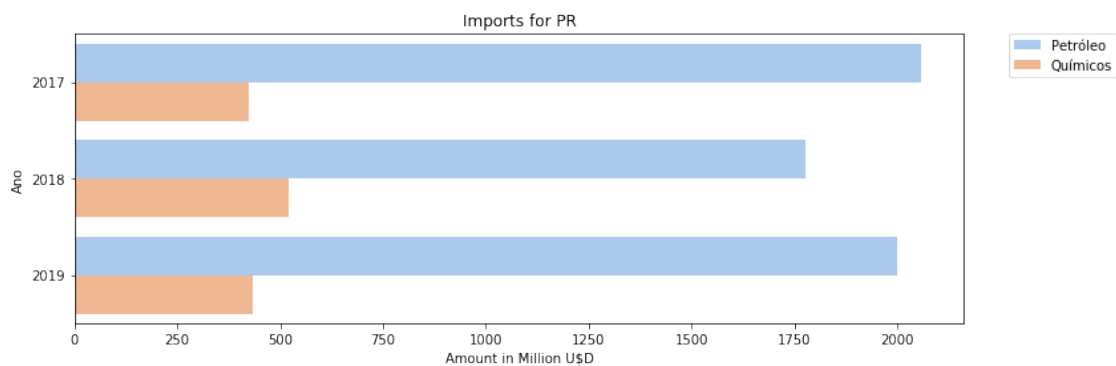


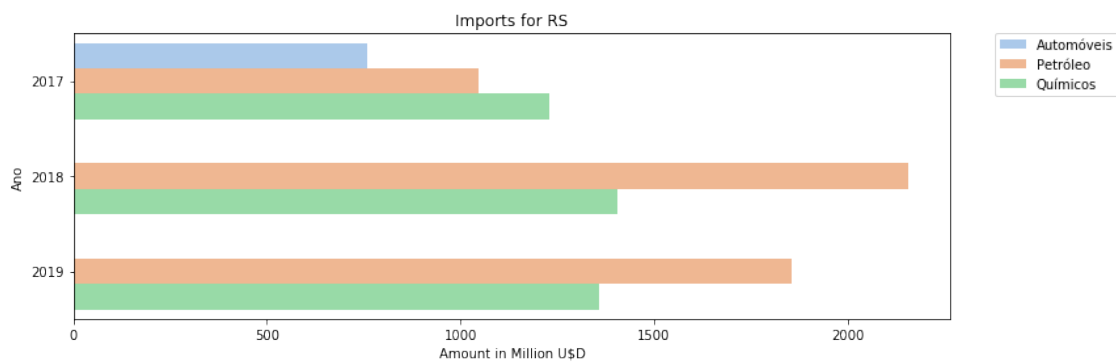
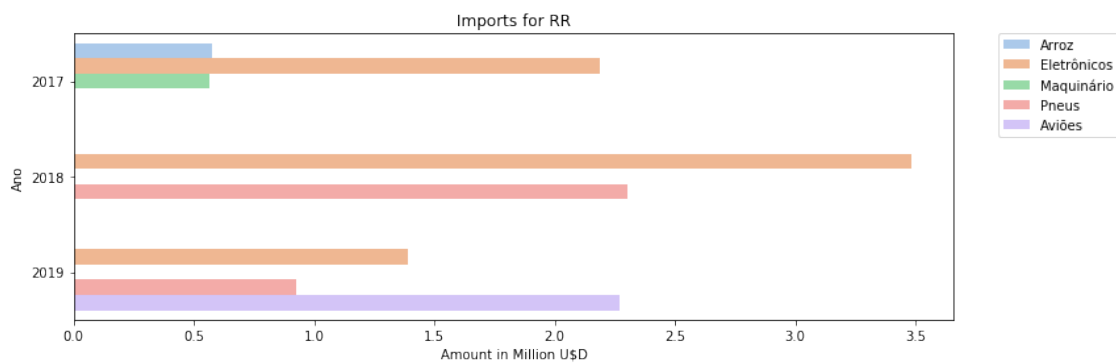
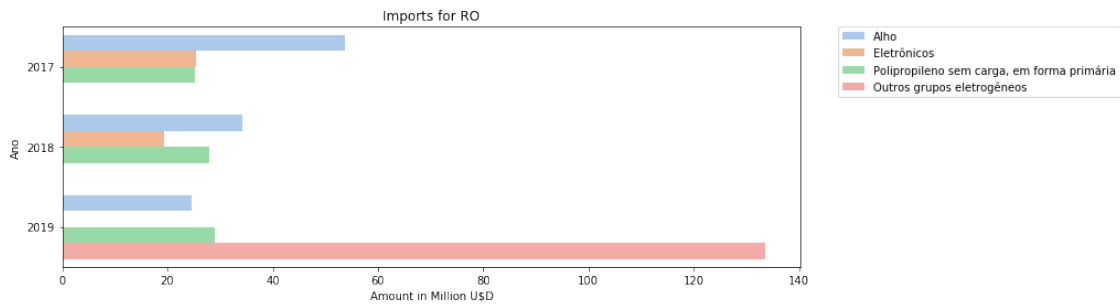


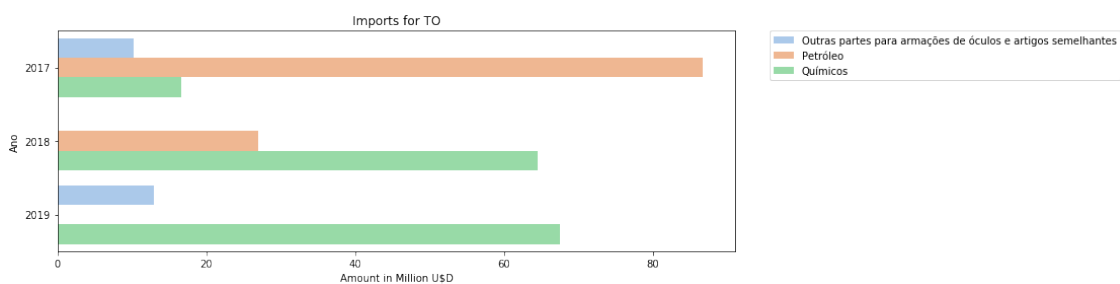
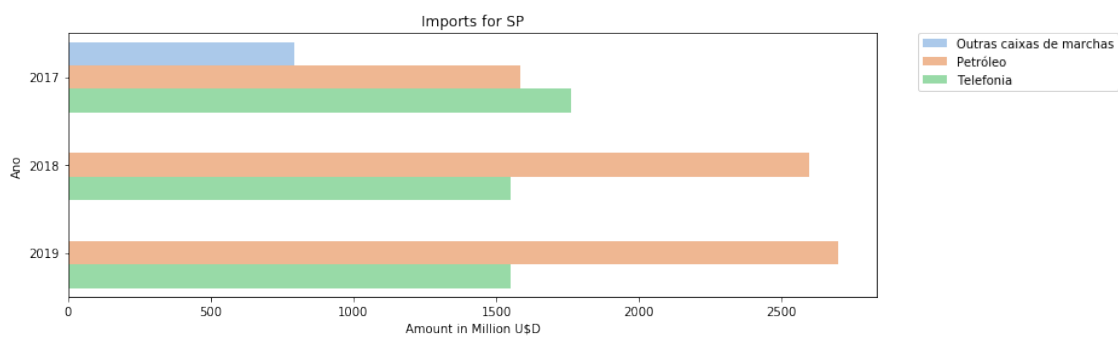
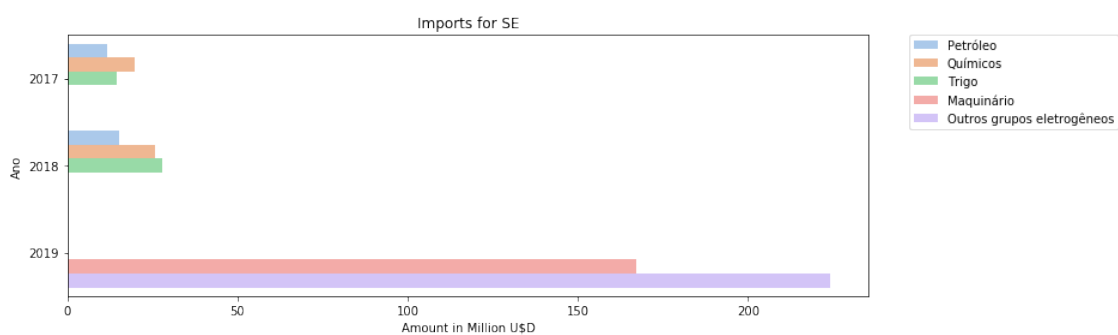
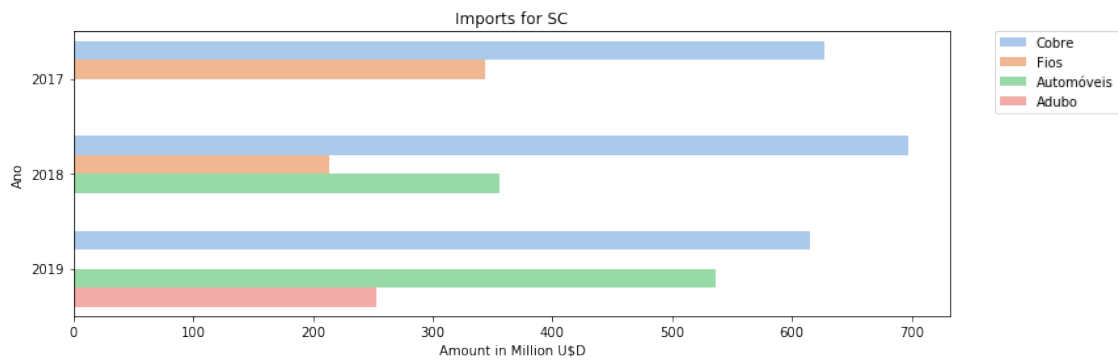




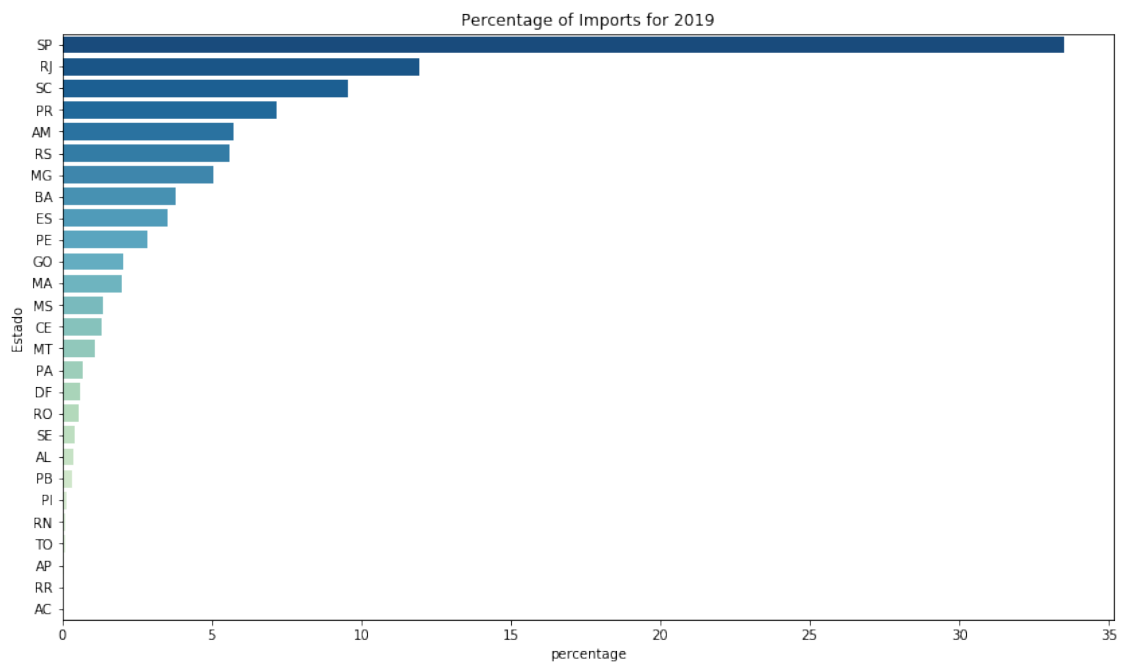








	Ano	Estado	Amount(USD)
0	2019	AC	1957132
1	2019	AL	665943197
2	2019	AM	10163473885
3	2019	AP	126462640
4	2019	BA	6772304966
5	2019	CE	2356778510
6	2019	DF	1108620808
7	2019	ES	6274742006
8	2019	GO	3584324347
9	2019	MA	3551463014



1.6 Prediction of 3 months ahead: Exports

I have downloaded two extra indexes to help in the task of prediction that have been suggested as being components of a model for prediction of exports and imports: the inflation rate and the currency of Brazilian Real against US Dollars.

Sources: <https://br.investing.com/currencies/usd-brl-historical-data>
<https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9256-indice-nacional-de-precos-ao-consumidor-amplo.html?=&t=series-historicas>

Another important aspect is to make a fair distinction of seasonal products (such as some agricultural commodities) and non seasonal products. In order to do this, we will plot a bar plot of top 3 products for this decade.