

# Omdena - Milan Chapter Agrifoods

## AI for Sustainable agri-food systems: use of Satellite Imagery

### Exploratory analysis of fruit production in Italy 2020-2022

Author: Maria Fisher

Here we have analysed the major areas and crop produced in Italy. Crop dataset used in this study was downloaded from the Italian National Institute of Statistics (Istat). The objective is to get some insights about crop production and its respective areas in Italy and how climate change have affected crop production in the last 3 years.

ITERATOR	Territory	TIPO_DATOS	Data type	AGRI_MADRE	Type of crop	TIME	Select time	Value	Flag	Codes	Flags
0	ITD05	Venezia	PA_EXT	production area - hectares	KIWI	kiwi	2020	2020	33	NaN	NaN
1	ITD05	Venezia	PA_EXT	production area - hectares	KIWI	kiwi	2021	2021	35	NaN	NaN
2	ITD05	Venezia	PA_EXT	production area - hectares	KIWI	kiwi	2022	2022	37	NaN	NaN
3	ITE45	Frosinone	ART	total area - hectares	POMETES	pomegranates	2020	2020	11	NaN	NaN
4	ITE45	Frosinone	ART	total area - hectares	POMETES	pomegranates	2021	2021	12	NaN	NaN

## Pre-processing dataset

```

DATASET SHAPE: (23823, 5)

-----
FEATURE DATA TYPES:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 23823 entries, 0 to 23822
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  ---
 0   City        23823 non-null  object
 1   Data_type   23823 non-null  object
 2   Type_crop   23823 non-null  object
 3   Year        23823 non-null  int64
 4   Value       23823 non-null  int64
dtypes: int64(2), object(3)
memory usage: 930.7+ KB
None

```

NUMBER OF UNIQUE VALUES PER FEATURE:				
City	107			
Data_type	4			
Type_crop	58			
Year	3			
Value	3119			
dtype: int64				

```

NULL VALUES PER FEATURE
City      0
Data_type 0
Type_crop 0
Year      0
Value     0
dtype: int64

```

23823 rows x 5 columns

DATASET SHAPE: (23823, 5)

```
FEATURE DATA TYPES:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 23823 entries, 0 to 23822
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  --
 0   City        23823 non-null    object
 1   Data_type   23823 non-null    object
 2   Type_crop   23823 non-null    object
 3   Year        23823 non-null    int64
 4   Value       23823 non-null    int64
dtypes: int64(2), object(3)
memory usage: 930.7+ KB
None
```

NUMBER OF UNIQUE VALUES PER FEATURE:

City: 187  
Data\_type: 4  
Type\_crop: 58  
Year: 3  
Value: 1119  
dtype: int64

NULL VALUES PER FEATURE

City: 0  
Data\_type: 0  
Type\_crop: 0  
Year: 0  
Value: 0  
dtype: int64

## Cities in Italy producing fruits

```
array(['production area - hectares', 'total area - hectares',  
      'harvested production - quintals', 'total production - quintals'],  
      dtype=object)
```

## Select only Values for total fruits production - quintals

City	Data_type	Type_crop	Year	Value	
23	Perugia	total production - quintals	nectarine	2020	2000
24	Perugia	total production - quintals	nectarine	2021	2000
25	Perugia	total production - quintals	nectarine	2022	600
28	Lucca	total production - quintals	plum	2020	2090
29	Lucca	total production - quintals	plum	2021	2100

City	Data_type	Type_crop	Year	Total_production	
23	Perugia	total production - quintals	nectarine	2020	2000
24	Perugia	total production - quintals	nectarine	2021	2000
25	Perugia	total production - quintals	nectarine	2022	600
28	Lucca	total production - quintals	plum	2020	2090
29	Lucca	total production - quintals	plum	2021	2100

...	...	...	...	...	...
23783	Nuoro	total production - quintals	citrus fruit	2021	45070
23787	Cagliari	total production - quintals	citrus fruit	2021	5721
23791	Oristano	total production - quintals	citrus fruit	2021	114990
23795	Barietta-Andria-Trani	total production - quintals	citrus fruit	2021	3050
23799	Sud Sardegna	total production - quintals	citrus fruit	2021	353009

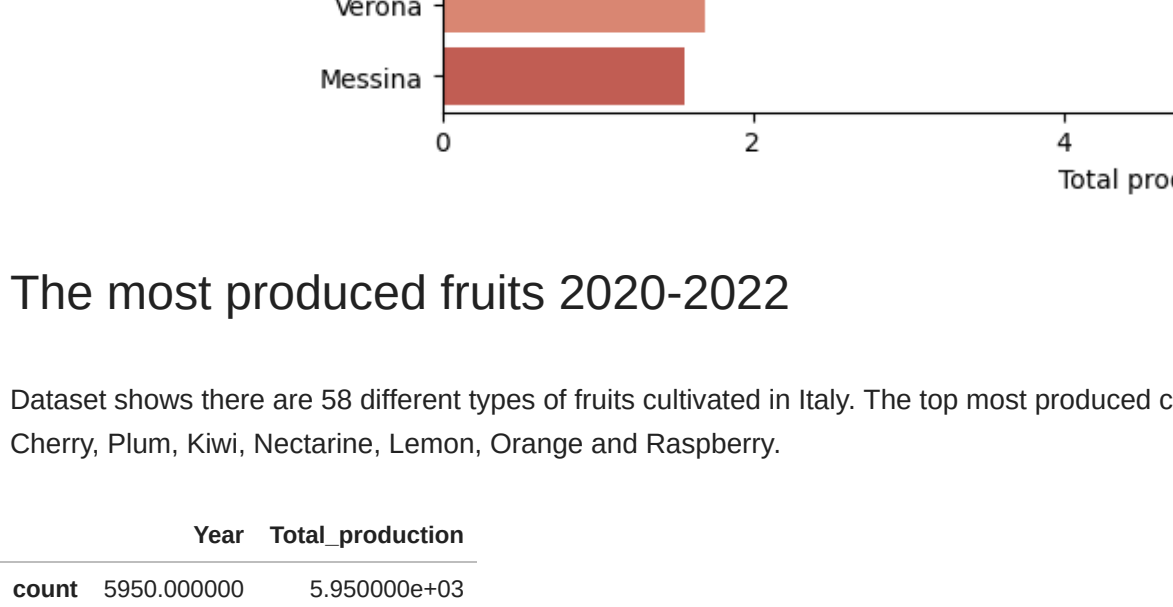
5950 rows x 5 columns

76	Reggio di Calabria	23195041
20	Caserta	22776411
103	Verona	16895734
53	Messina	15592550

Trentino Alto Adige / Sudtirolo	
Bolzano / Bozen	
Catania	
Siracusa	
Trento	

5950 rows x 4 columns

## Total fruit production in 2020-2022

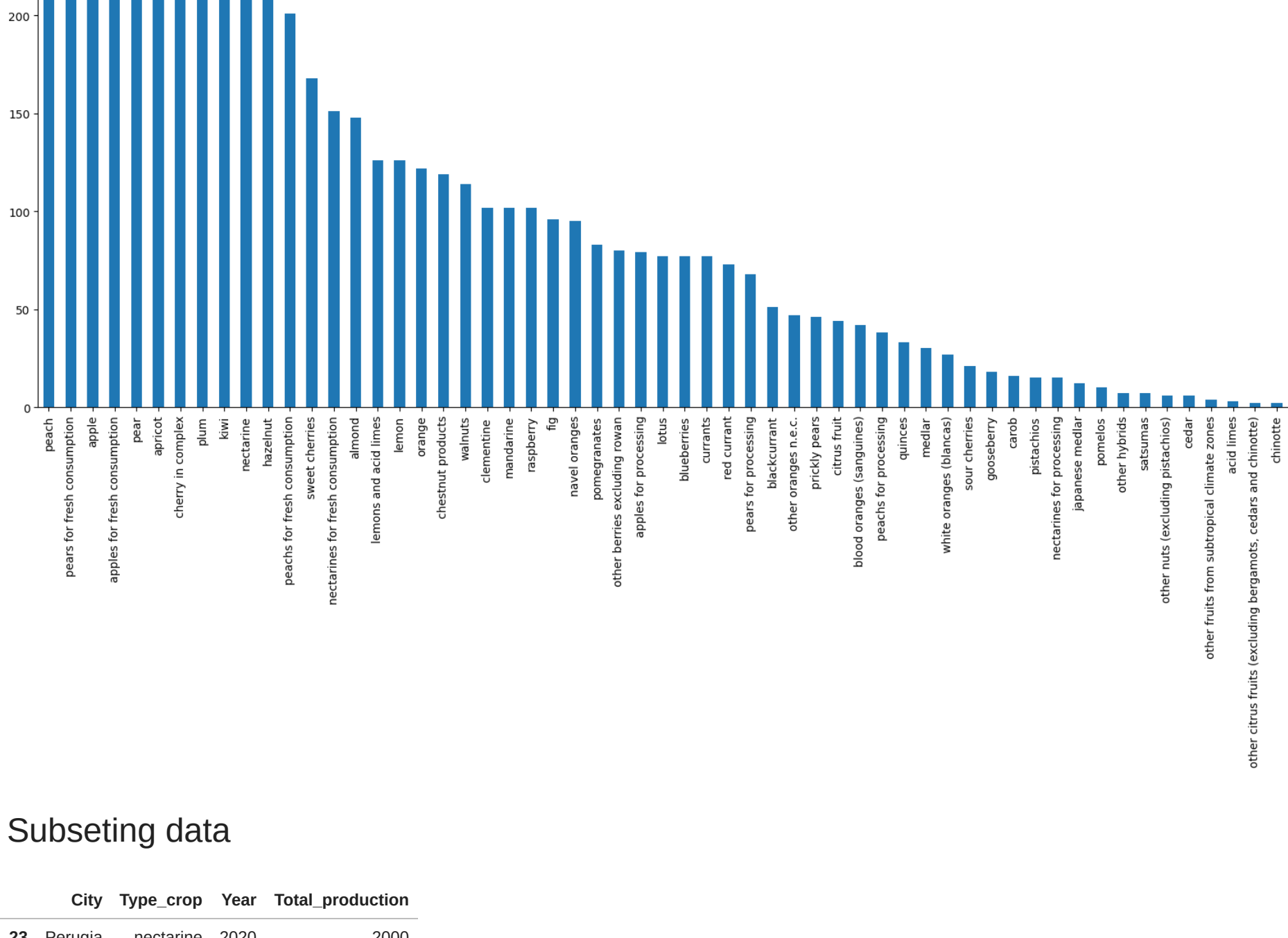


## Fruit production by Cities

25%	2020.000000	4.300000e+02
50%	2021.000000	2.775000e+03
75%	2022.000000	1.770000e+04
max	2022.000000	1.513440e+07

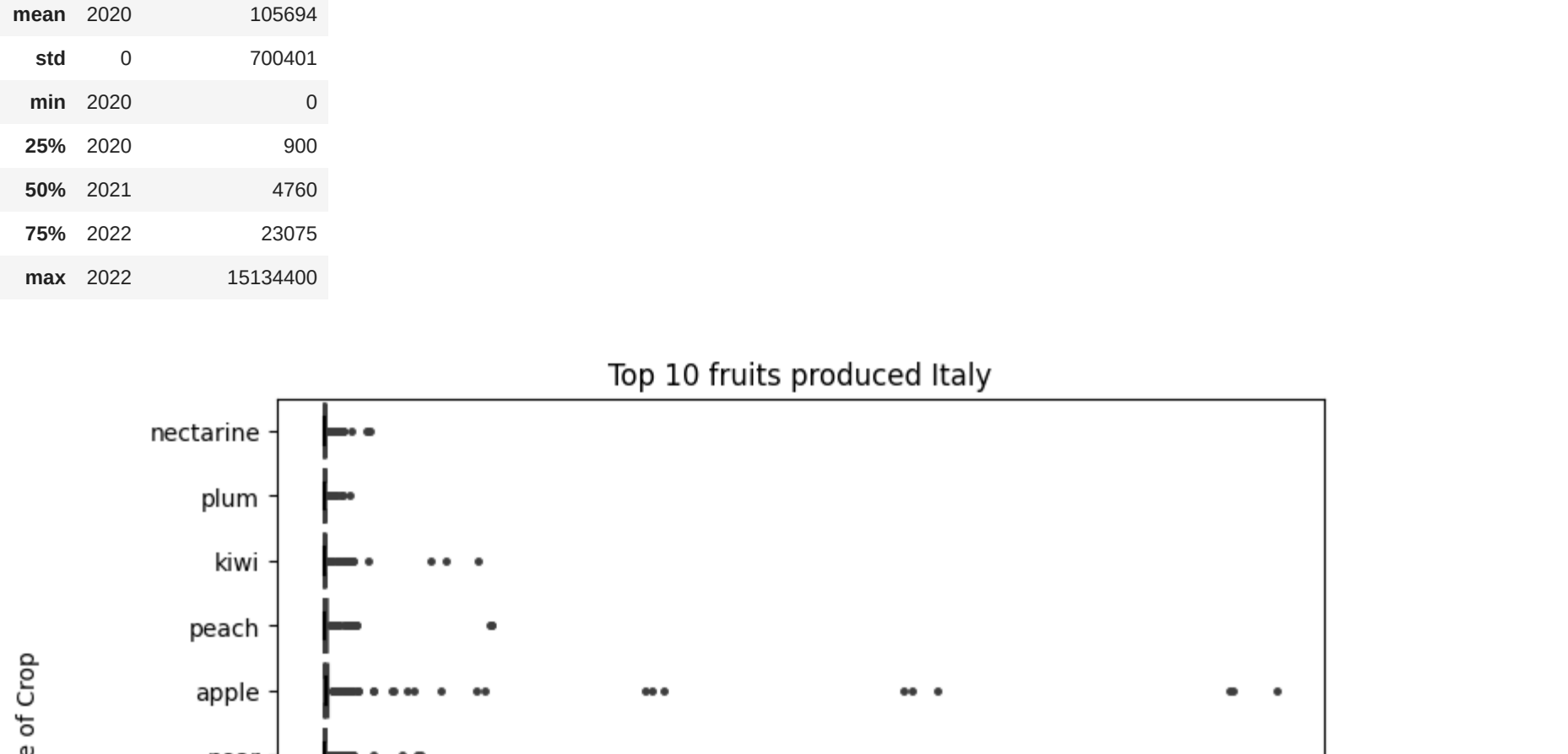
<AxesSubplot: >

Category	Value
1	300
2	295
3	290
4	285
5	280
6	275
7	270
8	265
9	260
10	255



## The most produced fruits 2020-2022

Dataset shows there are 58 different types of fruits cultivated in Italy. The top most produced crops are White oranges, Peach, Pear, Apples, Apricot, Cherry, Plum, Kiwi, Nectarine, Lemon, Orange and Raspberry.



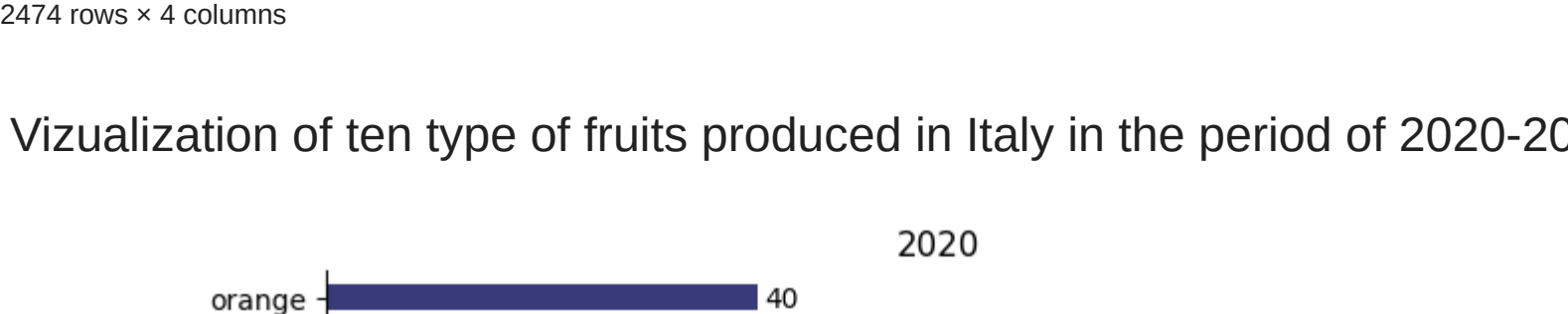
## Subsetting data

City	Type_crop	Year	Total_production	
23	Perugia	nectarine	2020	2000
24	Perugia	nectarine	2021	2000
25	Perugia	nectarine	2022	600
28	Lucca	plum	2020	2090
29	Lucca	plum	2021	2100

## Stats

Year	Total_production
count	2474
mean	2020
std	0
min	2020
25%	2020
50%	2021
75%	2022
max	2022

## Top 10 fruits produced Italy

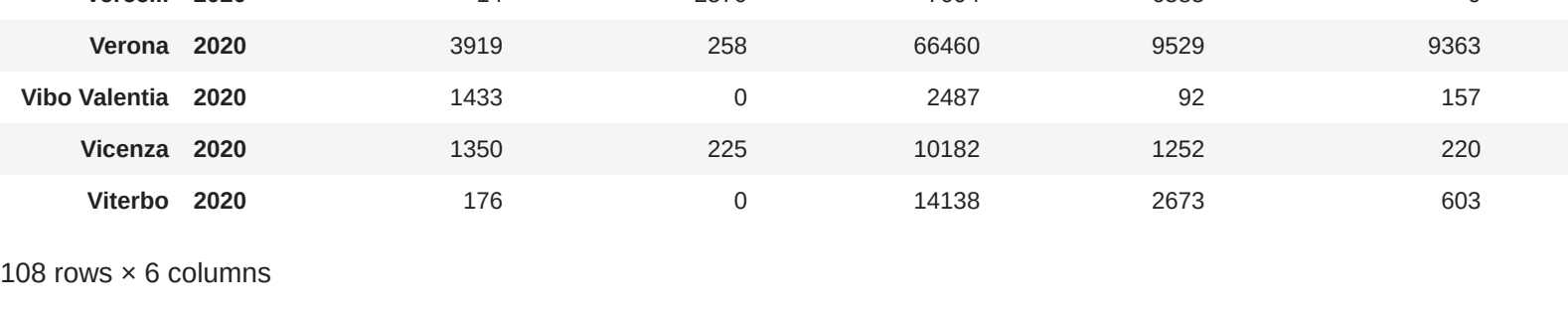
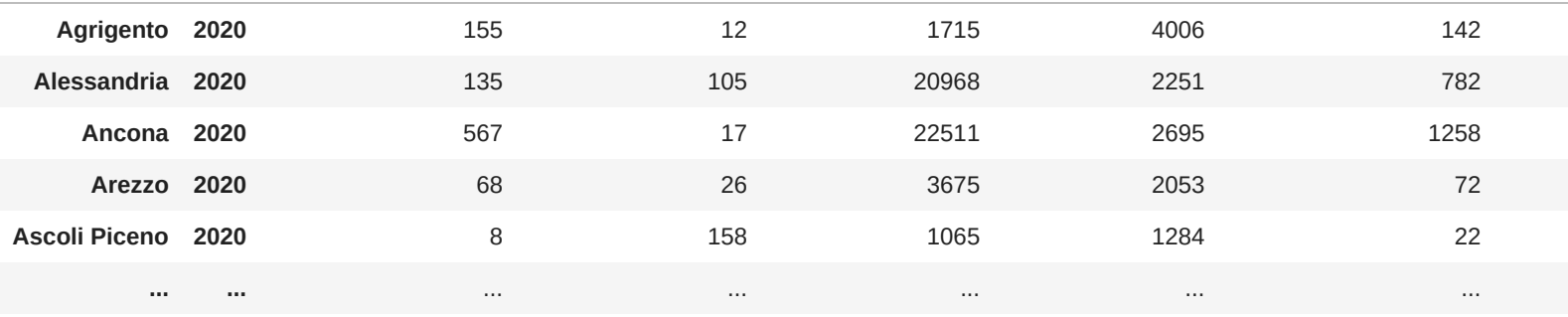
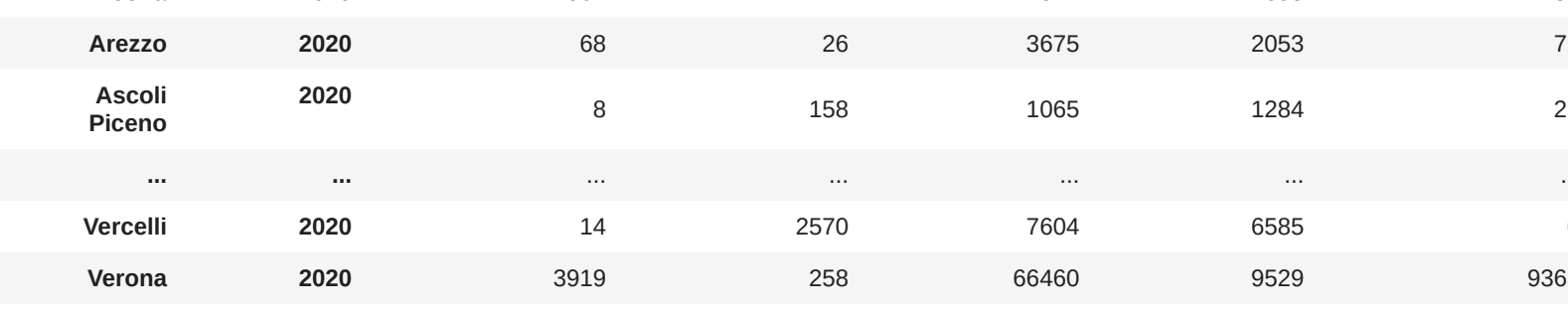
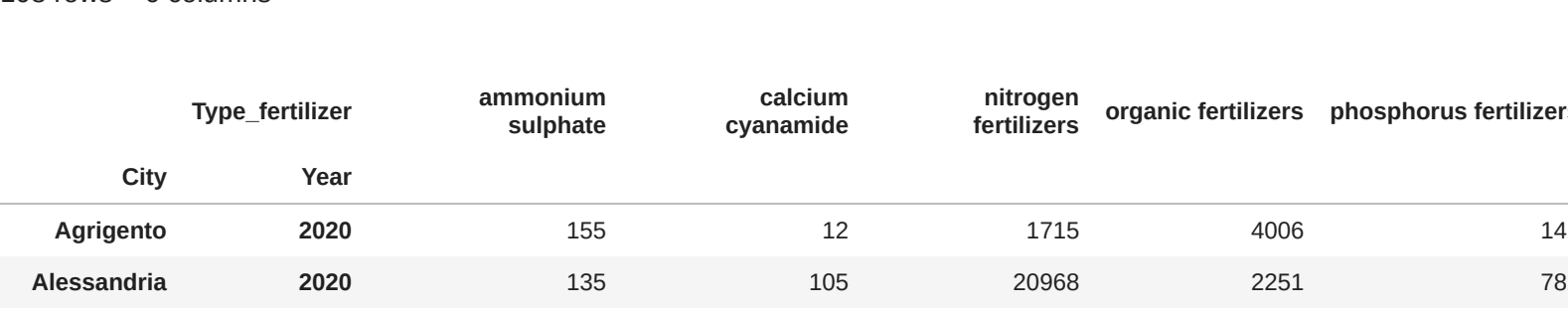
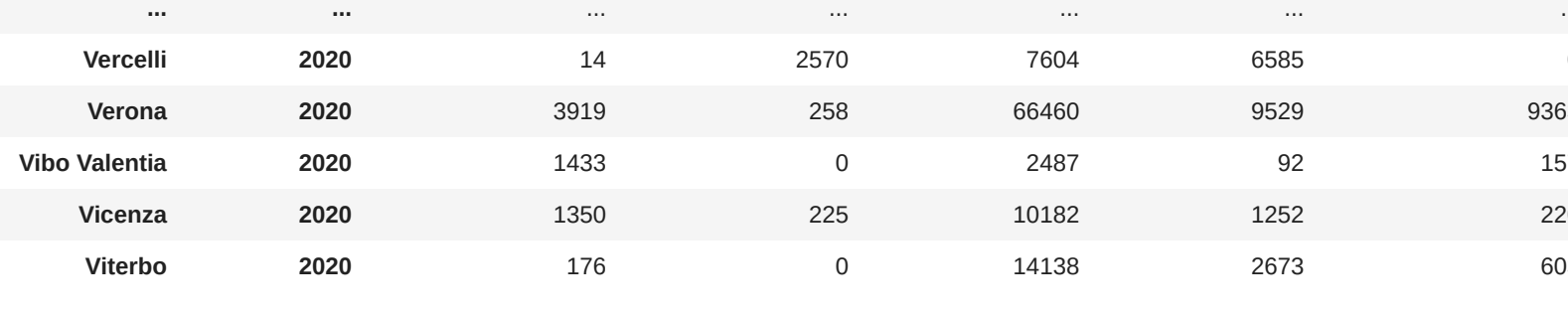
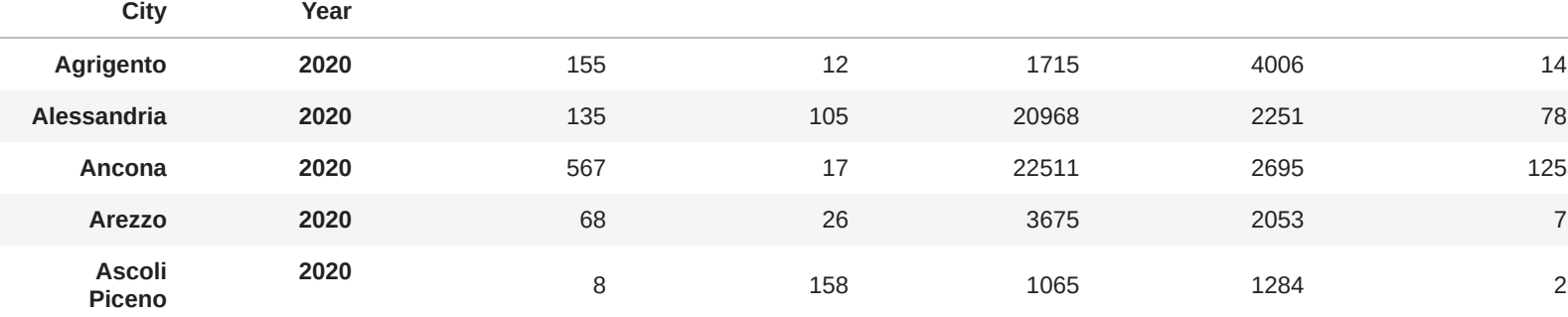
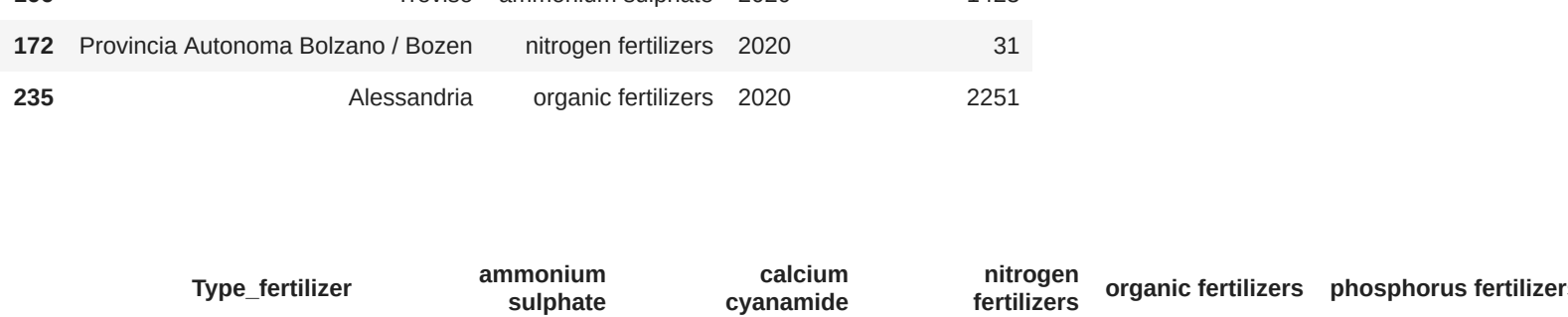
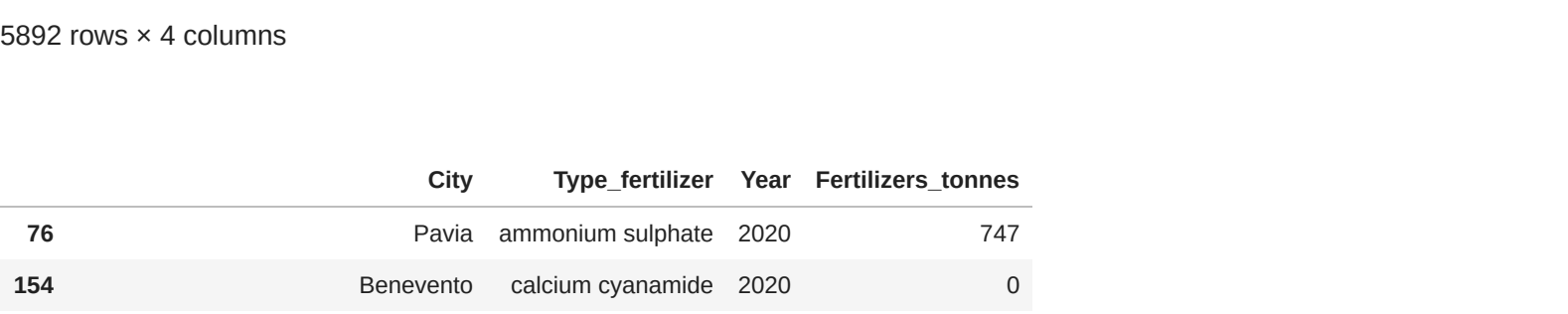
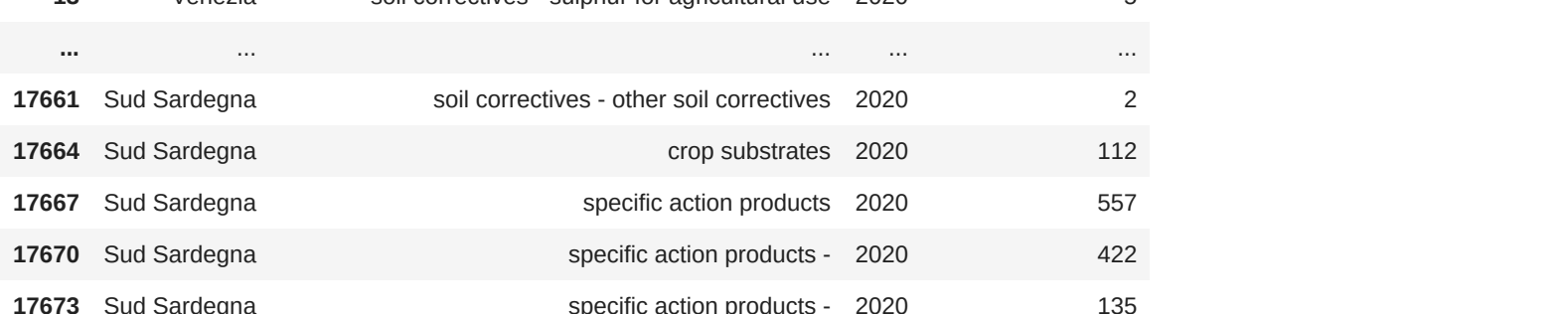
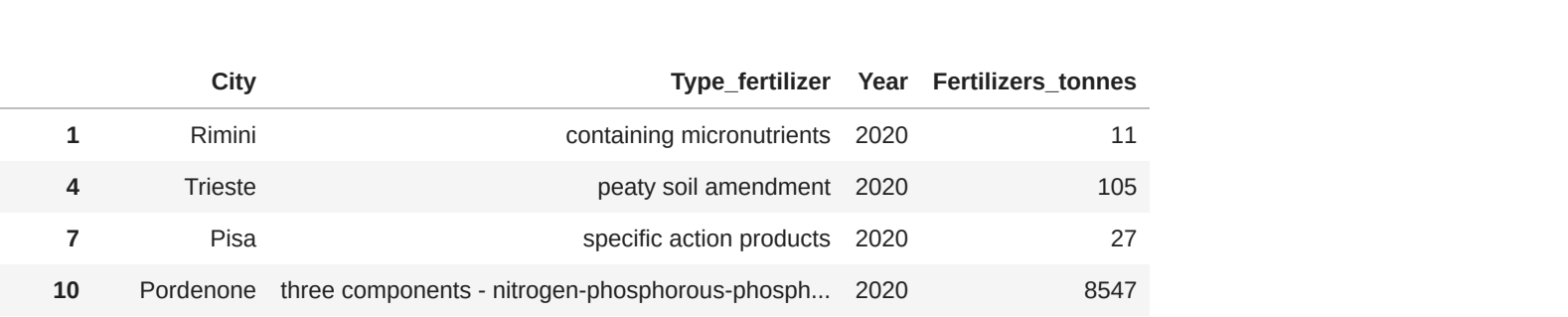
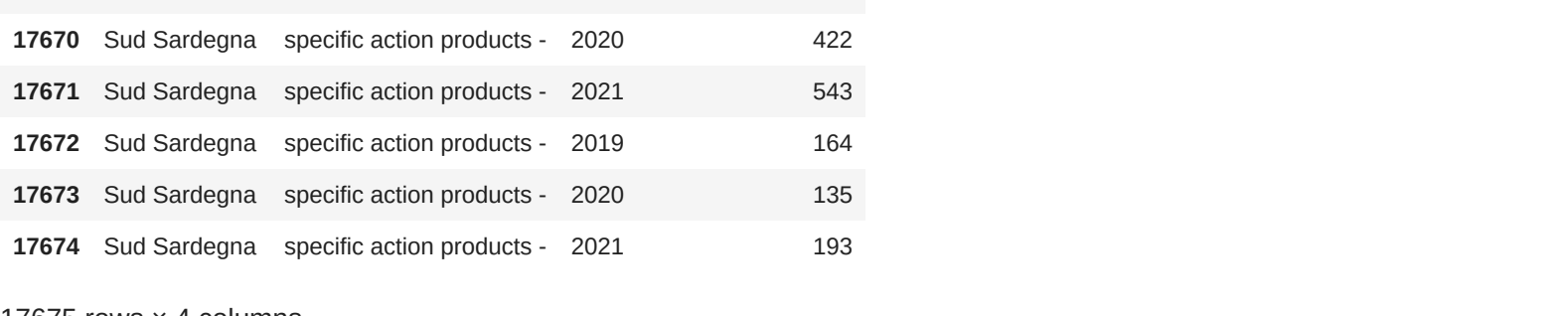
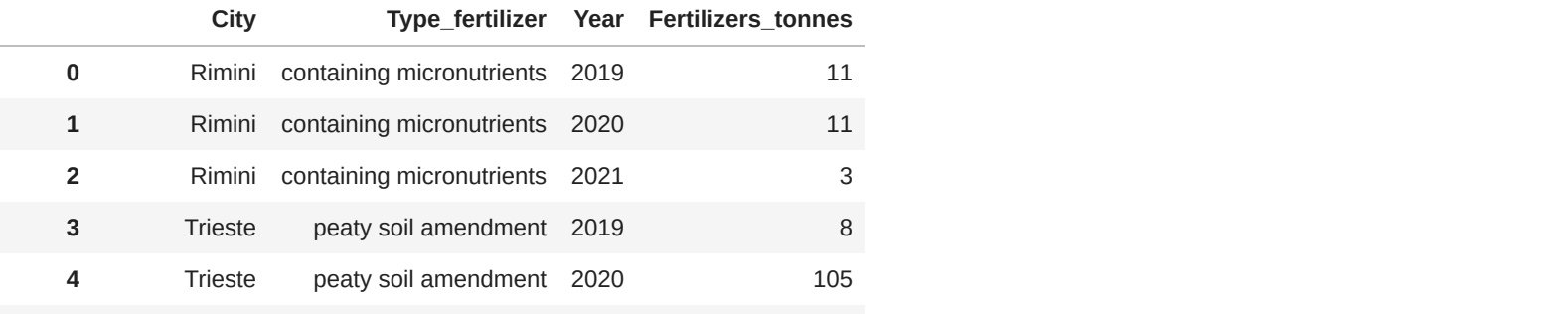
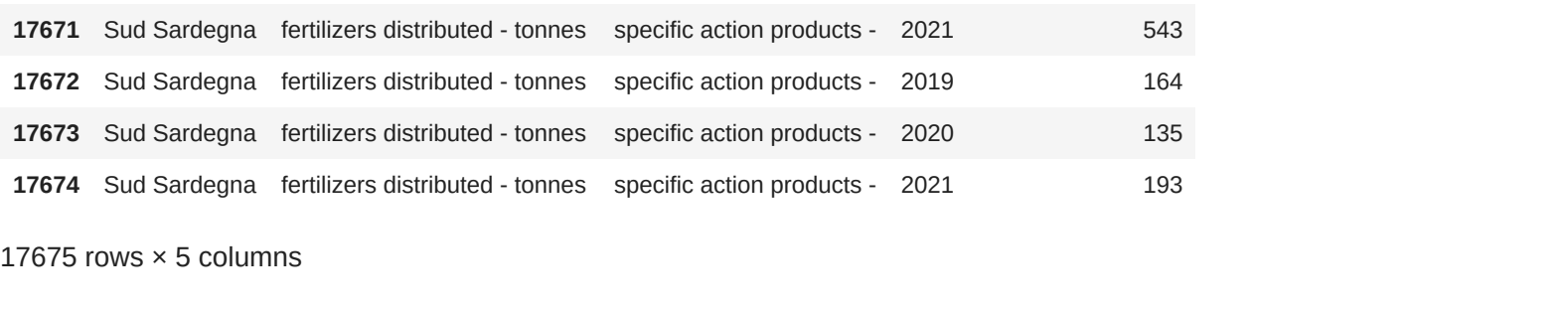
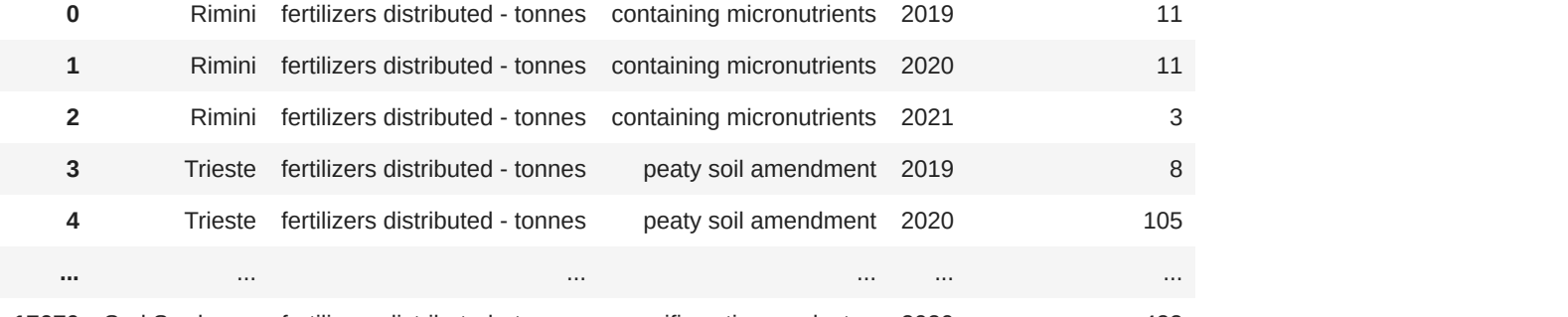
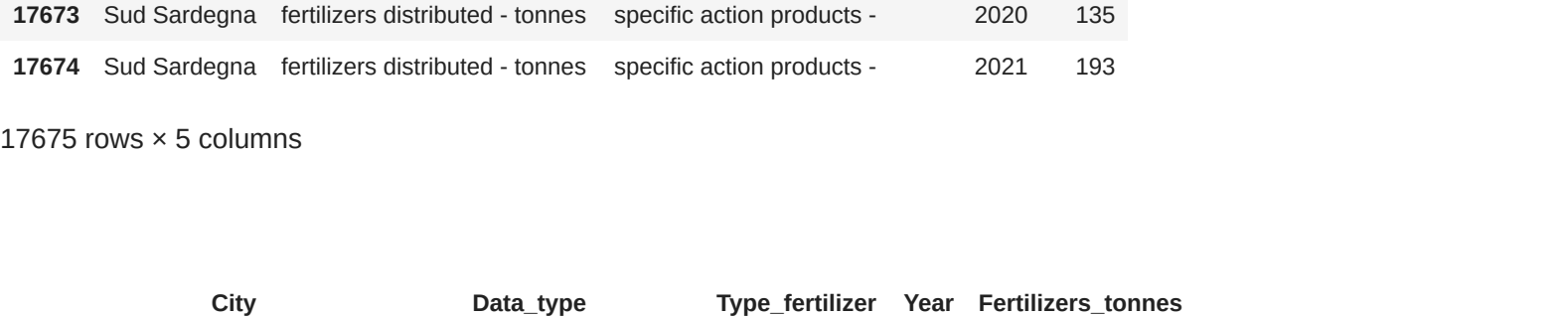
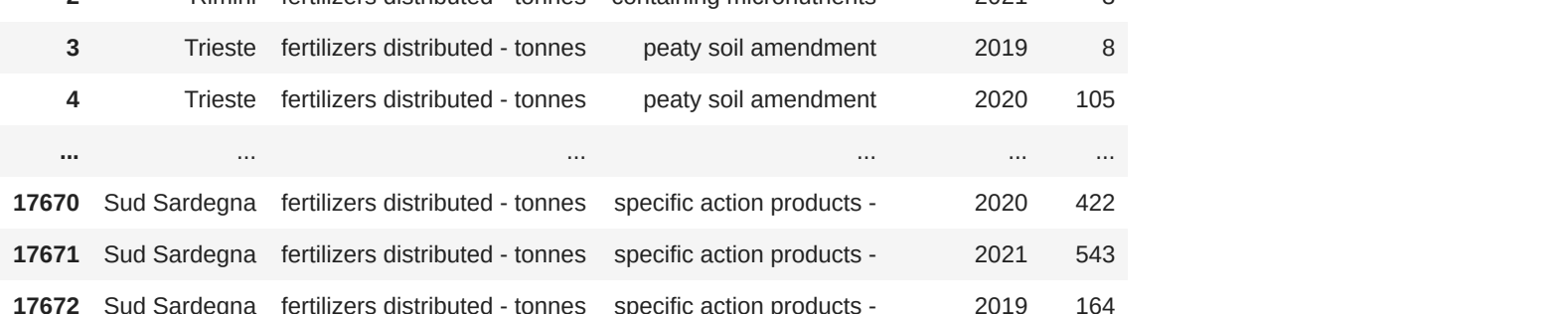
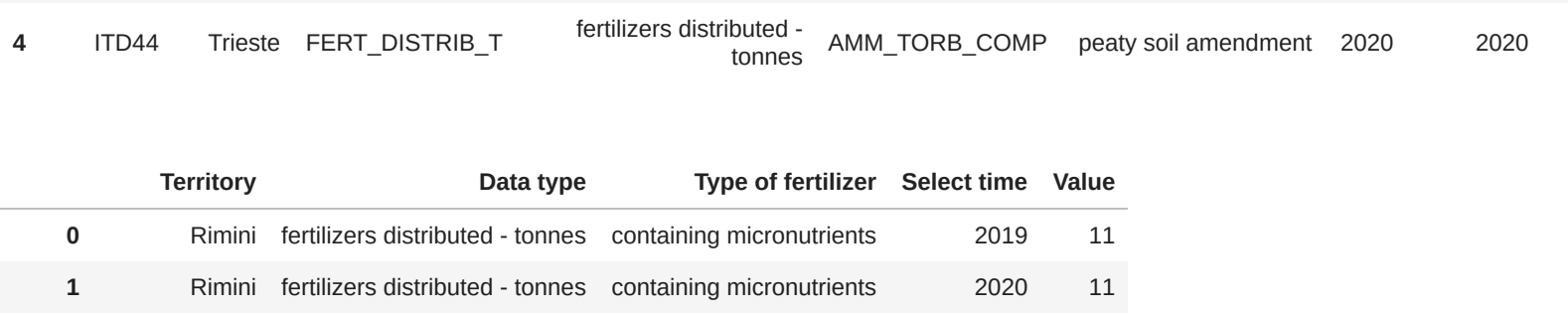
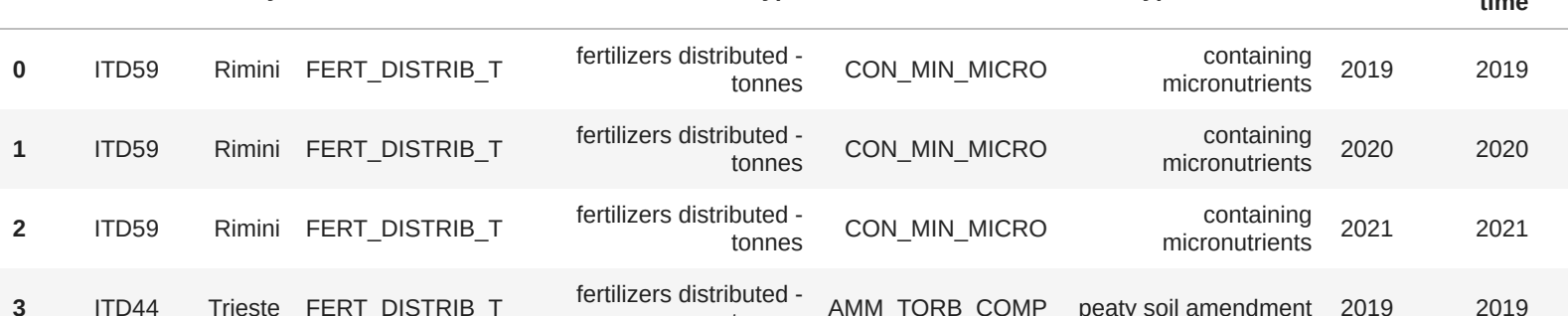
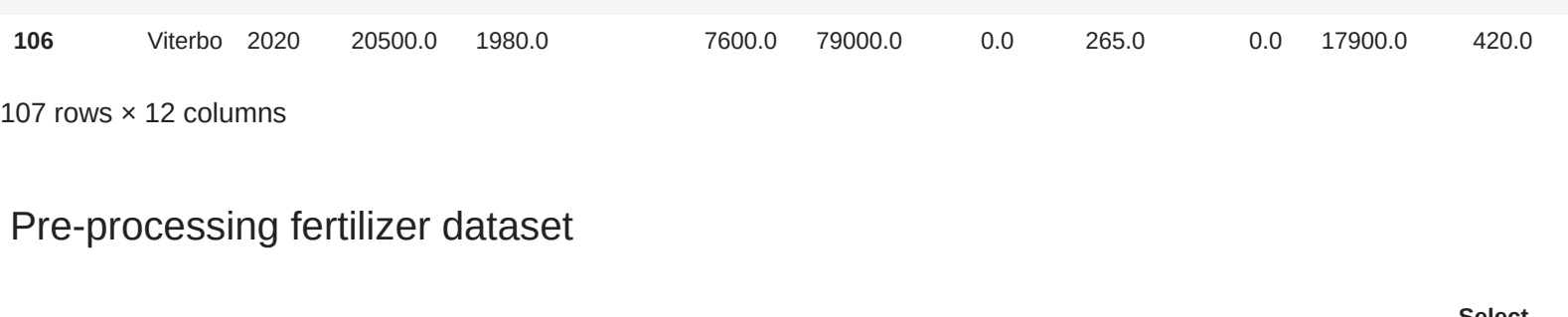
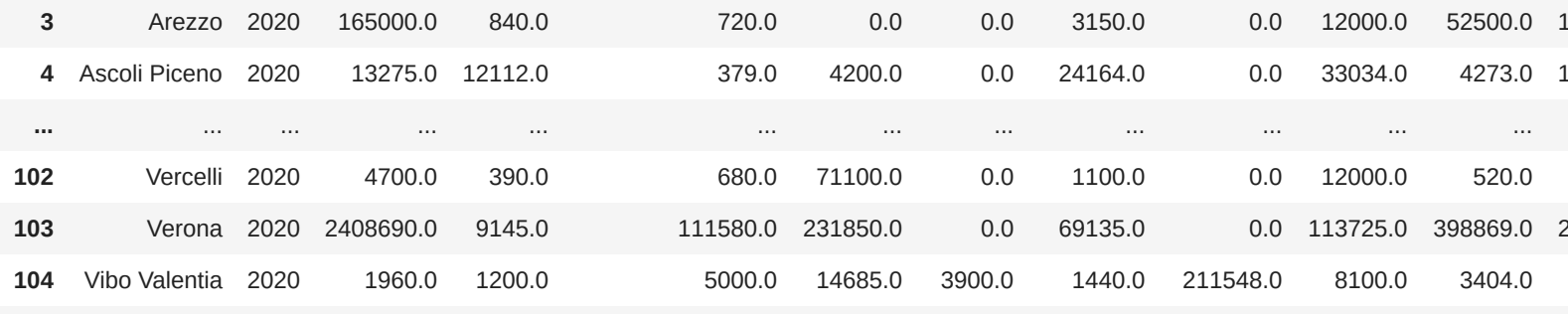
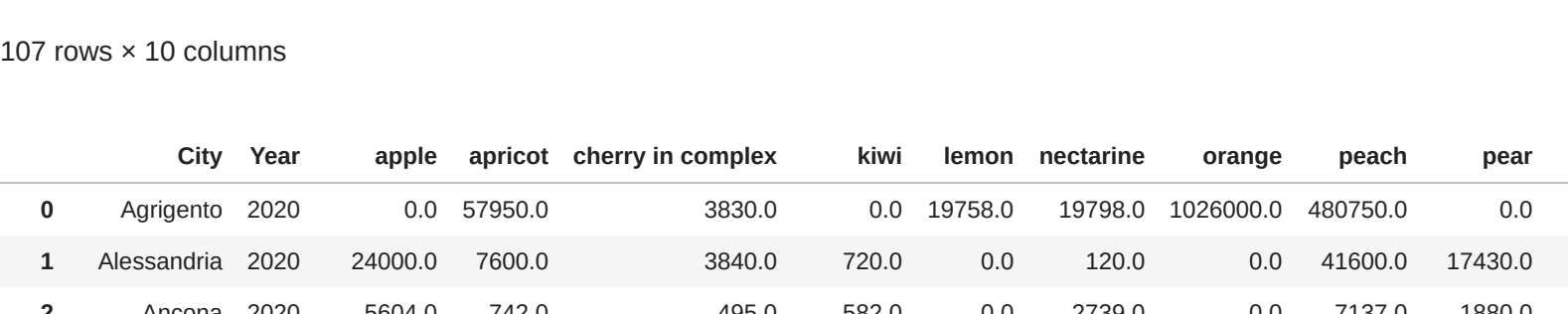
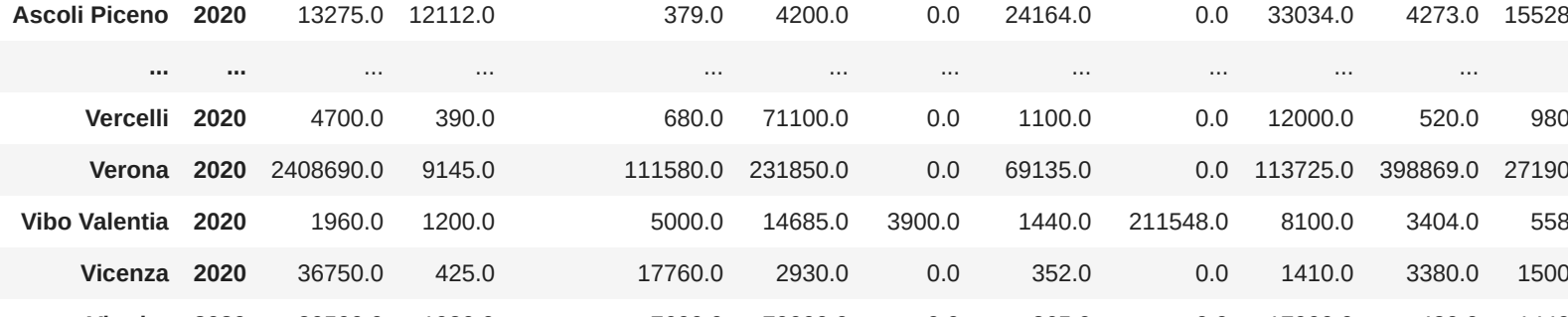
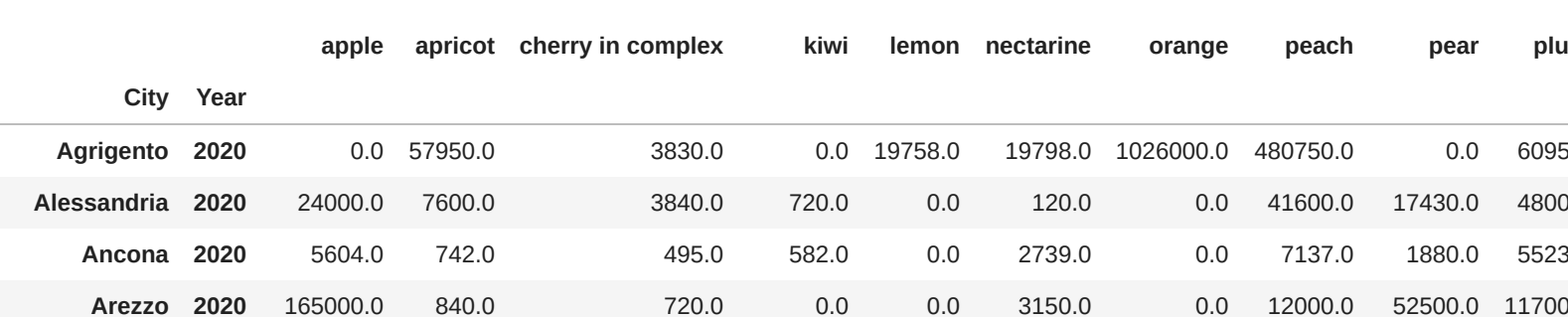
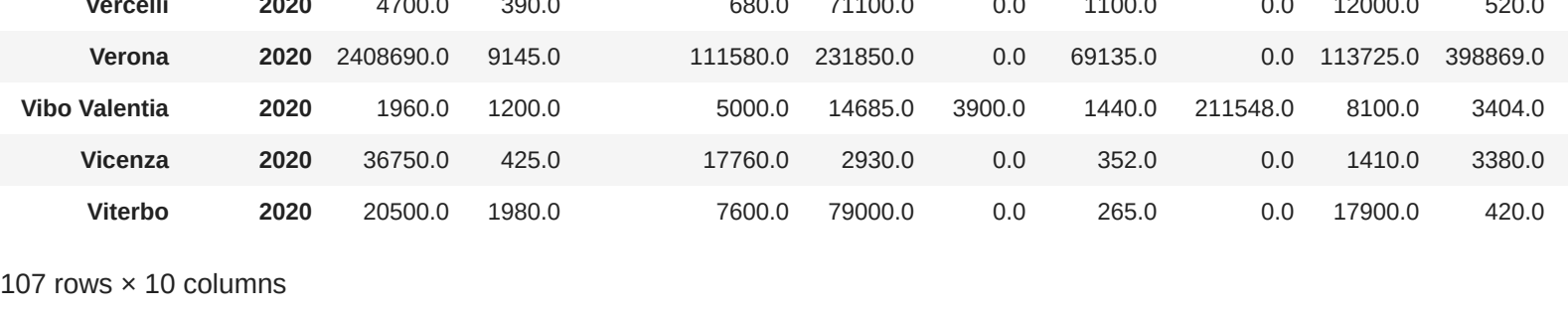
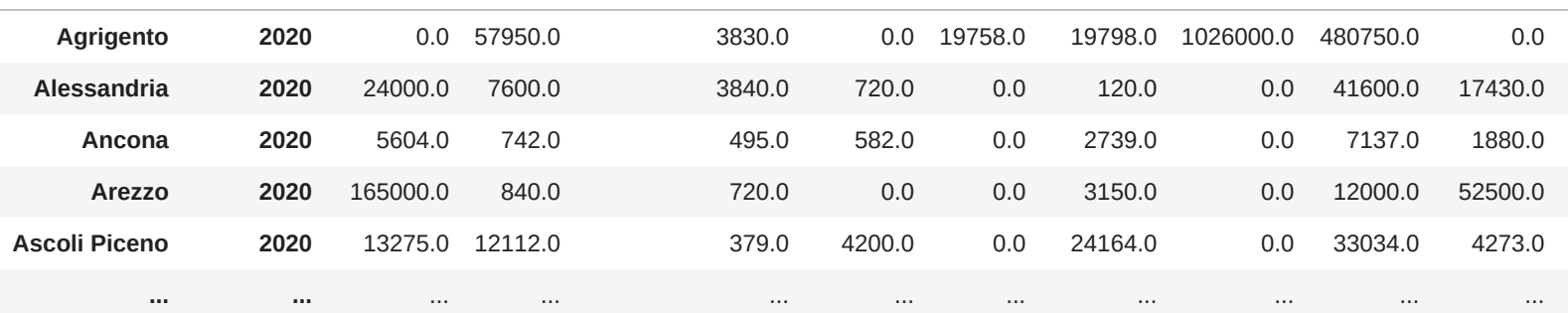
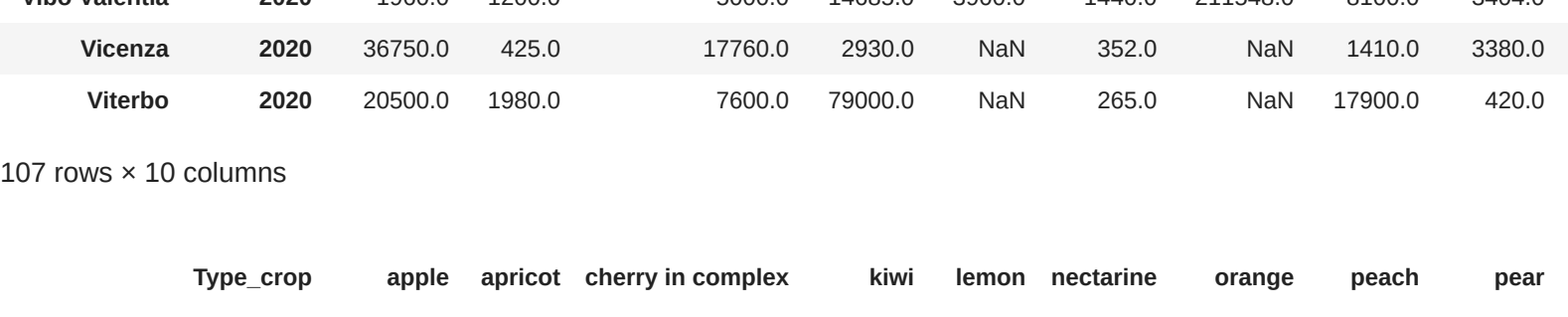
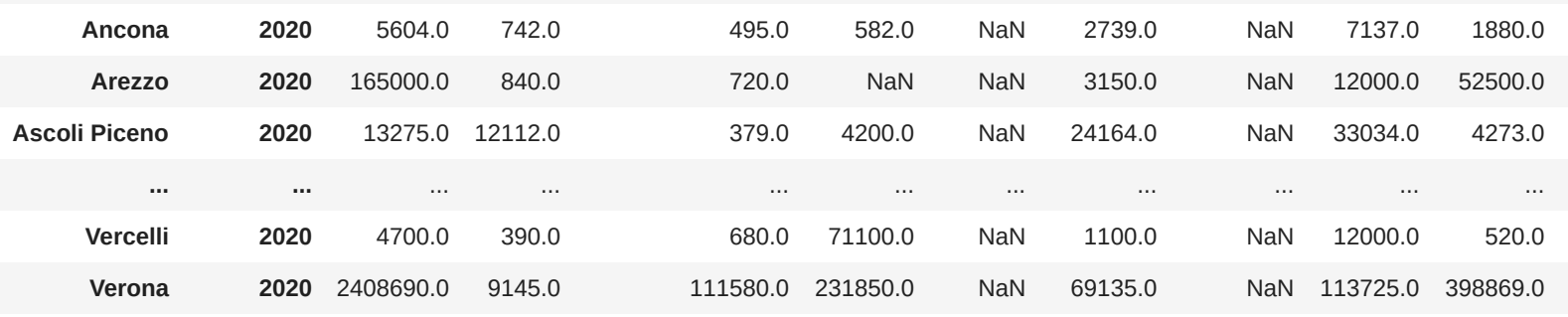
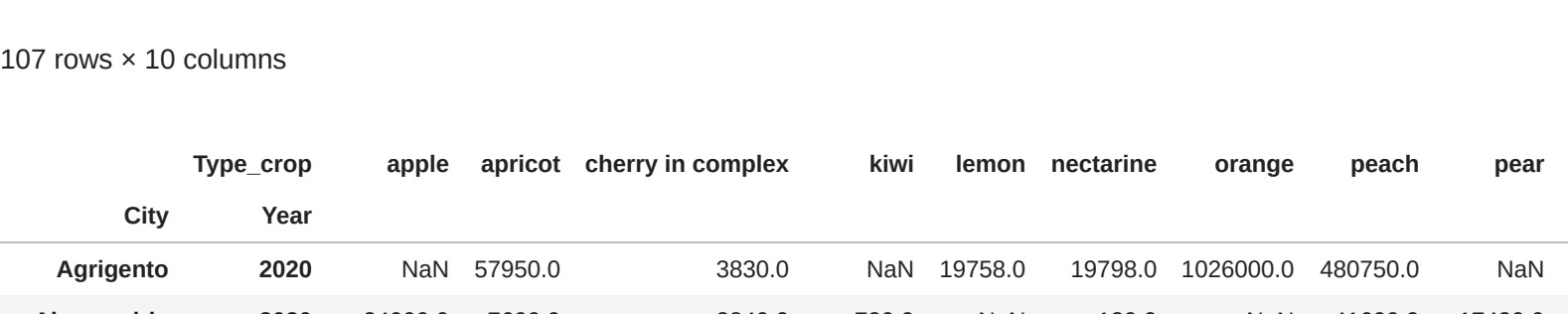
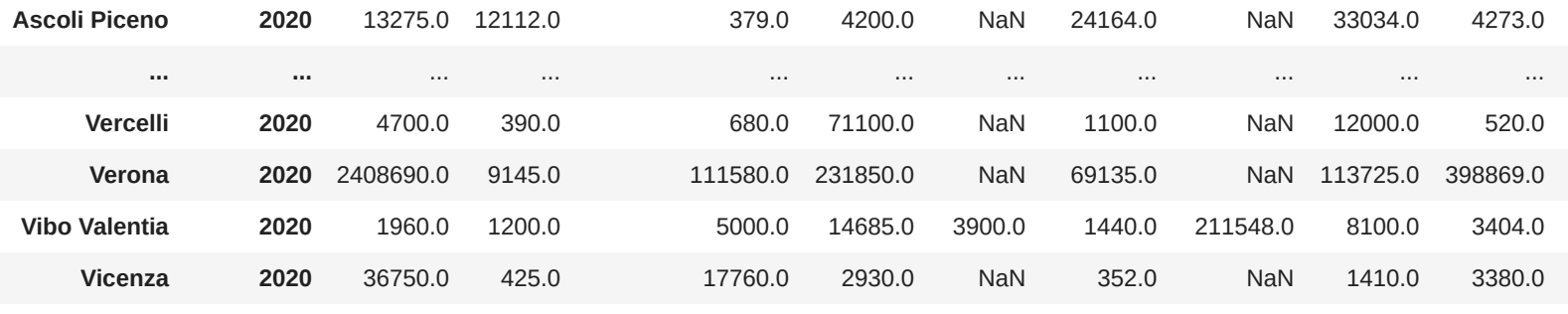
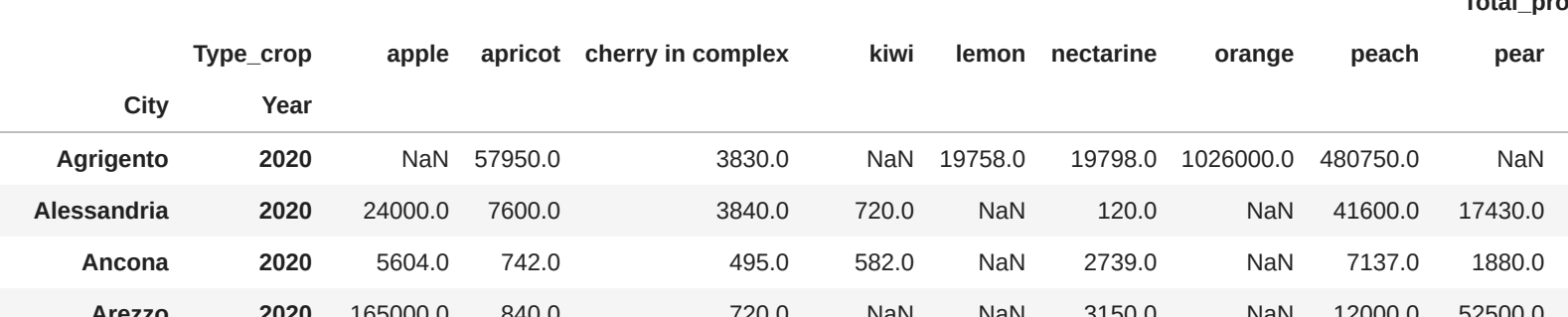
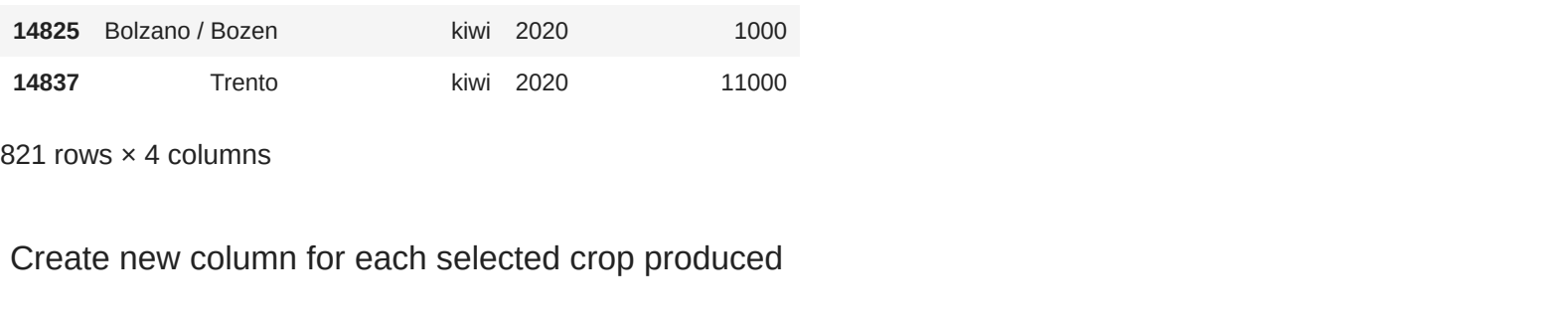
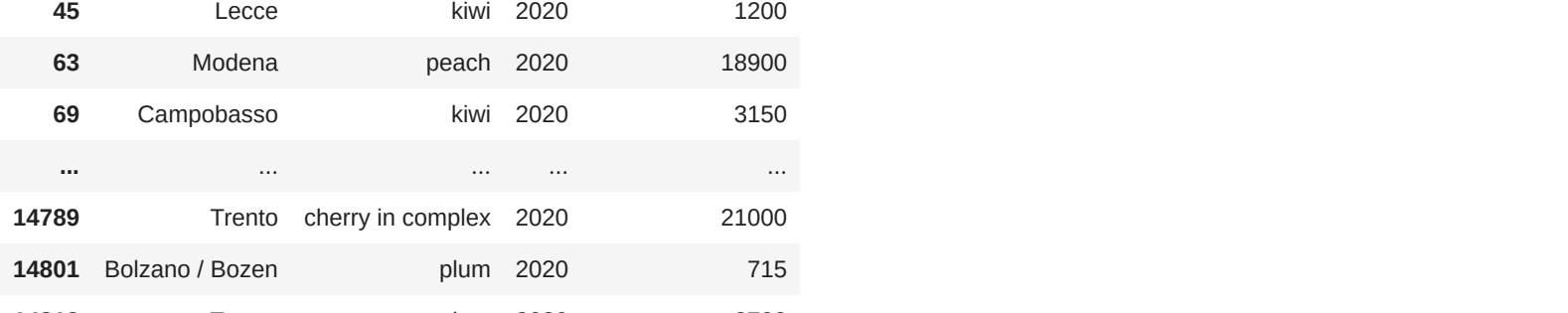
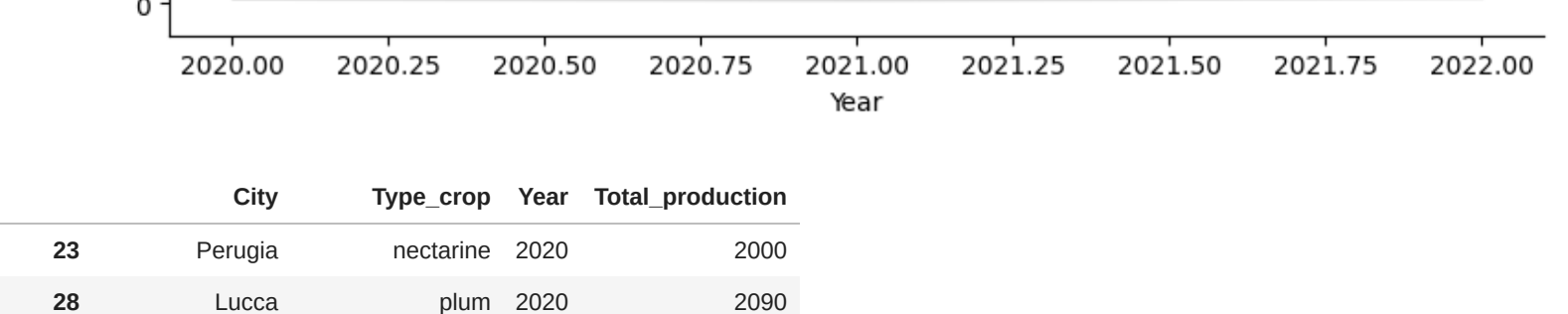
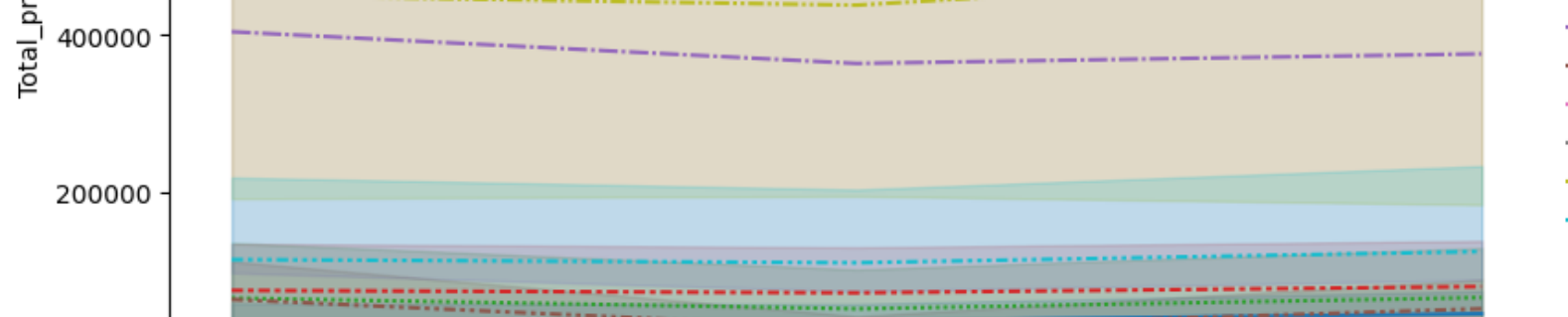
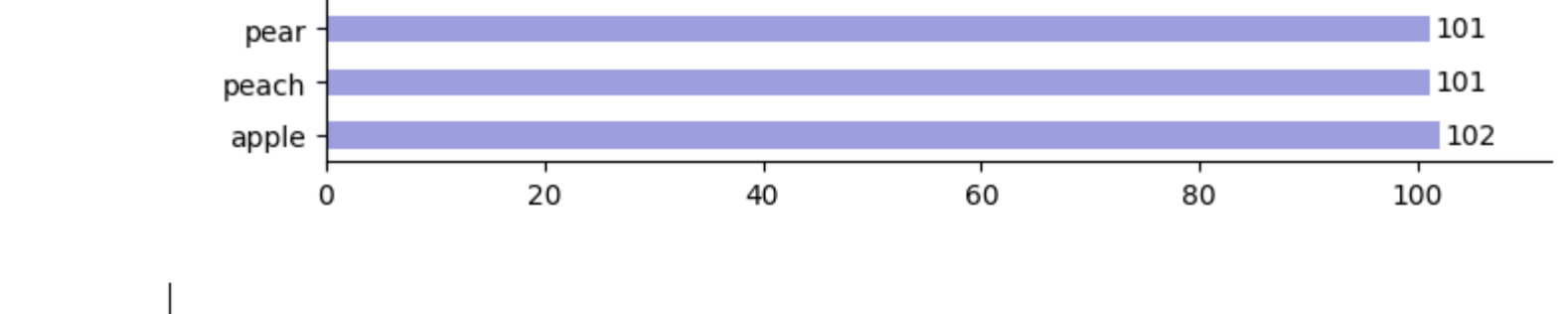
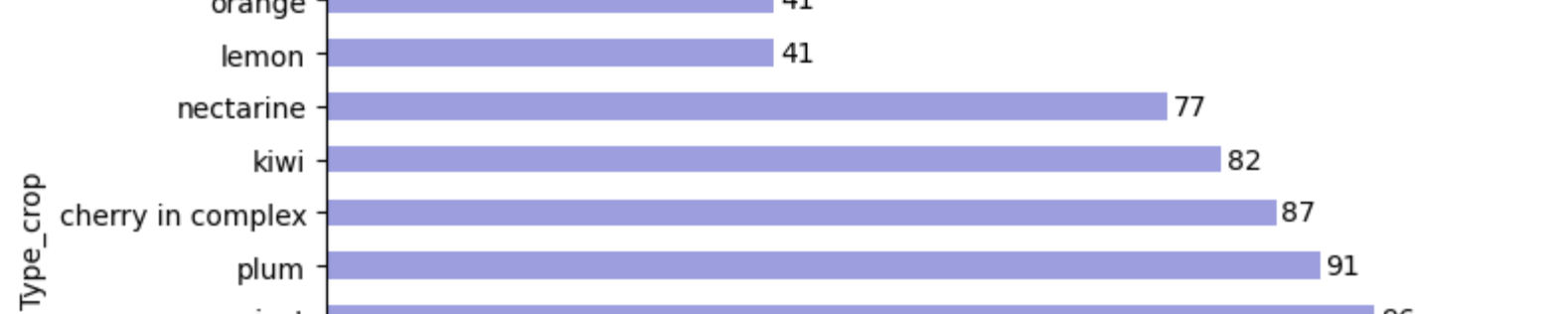
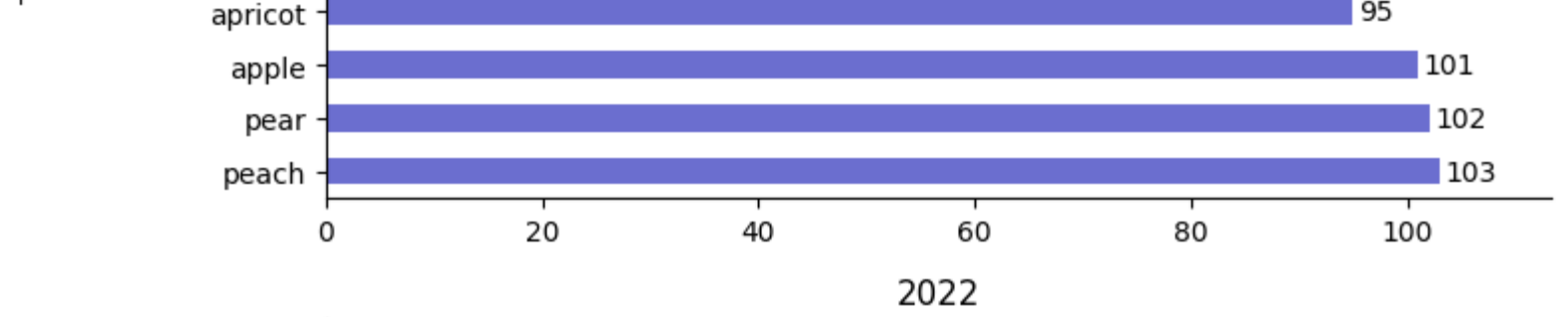


Type_crop	lemon
Type_crop	nectarine
Type_crop	kiwi
Type_crop	plum
Type_crop	cherry in complex
Type_crop	apricot

apple	0	20	40
peach	0	20	40
orange	0	20	40
lemon	0	20	40

2474 rows x 4 columns

## Vizualization of ten type of fruits produced in Italy in the period of 2020-2022

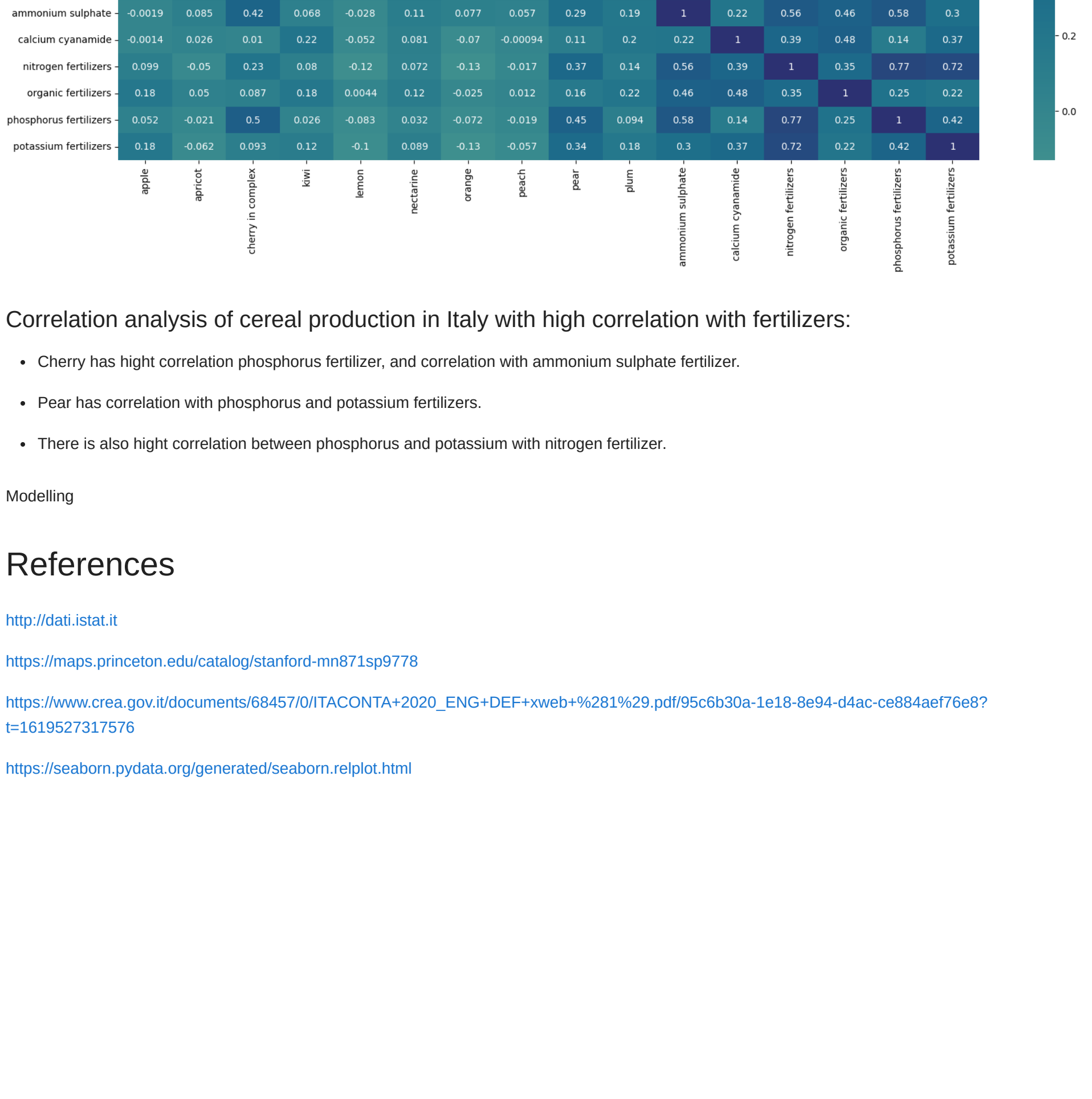




	City	Year	apple	apricot	cherry in complex	kiwi	lemon	nectarine	orange	peach	pear	plum	ammonium sulphate	calcium cyanamide	nit fert
0	Agrigento	2020	0.0	57950.0	3830.0	0.0	19758.0	19798.0	1026000.0	480750.0	0.0	6095.0	155.0	12.0	1
1	Alessandria	2020	24000.0	7690.0	3840.0	720.0	0.0	120.0	0.0	41600.0	17430.0	4800.0	135.0	105.0	2
2	Ancona	2020	5604.0	742.0	495.0	582.0	0.0	2739.0	0.0	7137.0	1880.0	5523.0	567.0	17.0	2
3	Arezzo	2020	165000.0	840.0	720.0	0.0	0.0	3150.0	0.0	12000.0	52500.0	11700.0	68.0	26.0	3
4	Ascoli Piceno	2020	13275.0	12112.0	379.0	4200.0	0.0	24164.0	0.0	33034.0	4273.0	15528.0	8.0	158.0	3
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
102	Verona	2020	4700.0	390.0	680.0	71100.0	0.0	1100.0	0.0	12000.0	520.0	980.0	14.0	2570.0	1
103	Verona	2020	2408890.0	8145.0	111580.0	231850.0	0.0	68135.0	0.0	113725.0	398869.0	27190.0	3919.0	258.0	6
104	Vibo Valentia	2020	1360.0	1200.0	5000.0	14685.0	3900.0	1440.0	211548.0	8100.0	3404.0	958.0	1433.0	0.0	2
105	Vicenza	2020	36750.0	425.0	17760.0	2930.0	0.0	352.0	0.0	1410.0	3380.0	1500.0	1350.0	225.0	1
106	Viterbo	2020	20500.0	1980.0	7600.0	79000.0	0.0	285.0	0.0	17900.0	420.0	1440.0	176.0	0.0	1

107 rows x 18 columns

## Correlation



Correlation analysis of cereal production in Italy with high correlation with fertilizers:

- Cherry has high correlation phosphorus fertilizer, and correlation with ammonium sulphate fertilizer.
- Pear has correlation with phosphorus and potassium fertilizers.
- There is also high correlation between phosphorus and potassium with nitrogen fertilizer.

Modelling

## References

<http://dati.istat.it>

<https://maps.princeton.edu/catalog/stanford-mm871sp9778>

[https://www.crea.gov.it/documents/68457/0/ITACONTA+2020\\_ENG+DEF+web+%281%29.pdf/955cb30a-1e18-8e94-d4ac-ce884ae76e8?i=1619527317576](https://www.crea.gov.it/documents/68457/0/ITACONTA+2020_ENG+DEF+web+%281%29.pdf/955cb30a-1e18-8e94-d4ac-ce884ae76e8?i=1619527317576)

<https://seaborn.pydata.org/generated/seaborn.relplot.html>