Registered Unemployment in Barcelona

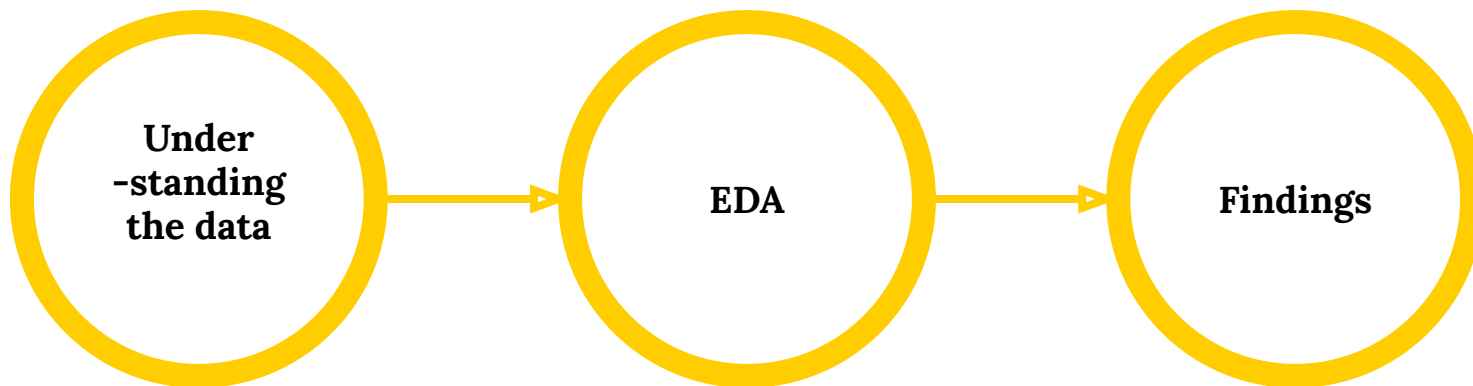# Hello!

I am **Vivian kim**

@https://www.linkedin.com/in/vivian-kim-89ab a9199/

Presentation itinerary

Under -standing the data → EDA → Findings

Source and Data

Registered unemployement in the city of Barcelona

⬛ Economy and Business   **⬛ Employment**

Registered unemployement in the city of Barcelona.

⬛ View complementary description and field definition

Open Data **BCN**

Souce: <u>here</u>

- ◎   **11 csv files from 2011 to 2021.**

- ◎   **Data regarding the registered unemployment in Barcelona per district/neighbourhood**

- ◎   **(Registered and people who applied for the registration)**

Data Cleaning

Combined 11 csv files to one dataframe

Translate Columns & Values

```
df2 = df2.rename(columns={'Any':'year',
                    'Mes':'month',
                    'Codi_Districte':'district_code',
                    'Nom_Districte':'district_name',
                    'Codi_Barri':'neighbourhood_code',
                    'Nom_Barri':'neighbourhood_name',
                    'Sexe':'sex','Demanda_ocupació':'Demand_occupation',
                    'Nombre':'number', })
df2.head(2)
```

	year	month	district_code	district_name	neighbourhood_code	neighbourhood_name	sex	Demand_occupation	number
0	2019	1	1	Ciutat Vella	1	el Raval	Homes	Registered Unemployment	1900
1	2019	1	1	Ciutat Vella	2	el Barri Gòtic	Homes	Registered Unemployment	485

Variable Selection and Merging

1. The total number of unemployed people (registered + applied for the registration per year) –>Q1
2. The total number of unemployed people per district –>Q2
3. The total number of unemployed women/men –>Q3
4. The total number of unemployed people (monthly data) –>Q4

Main Questions

1) Is there a trend?

And what caused the trend?

2)Which district has the highest unemployment rate ?

3) The difference between the total numbers of unemployment of men and women

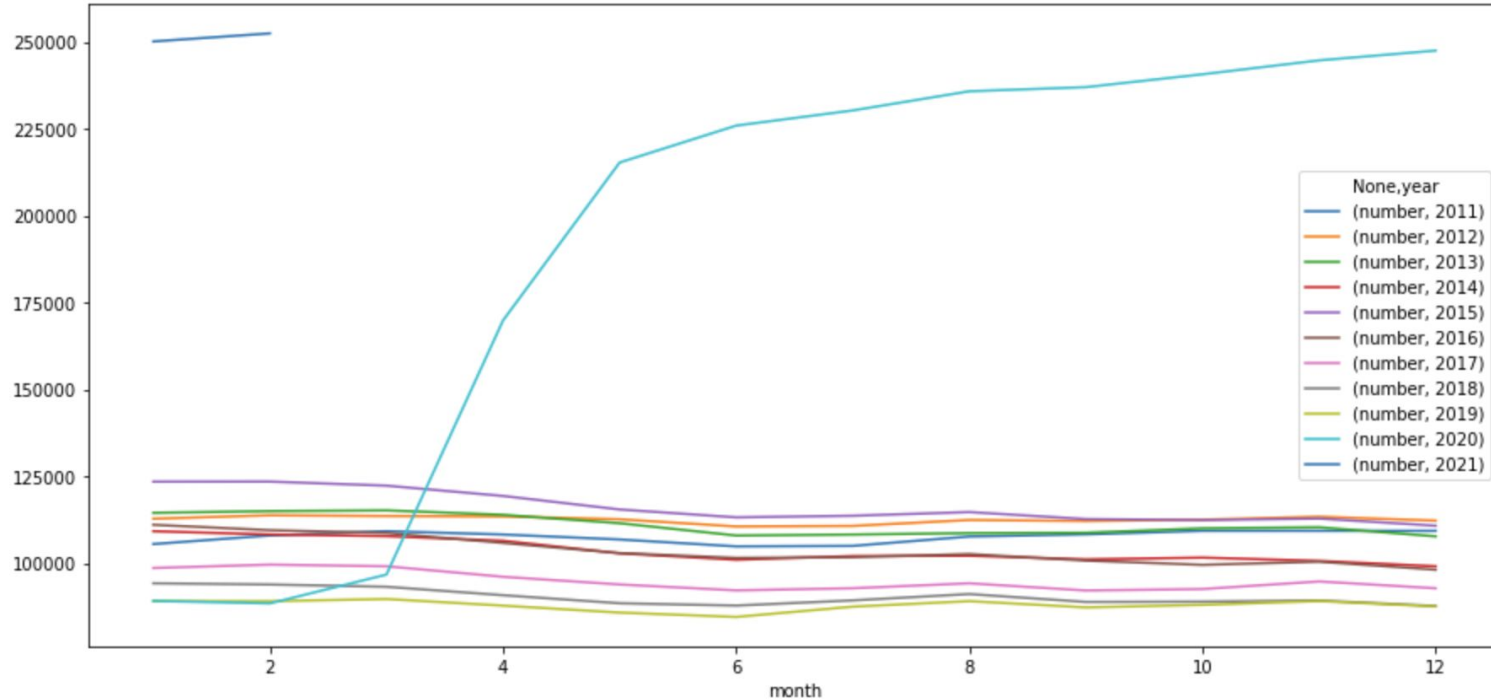3) Can we predict the next months unemployment rate?

EDA

1 Issue

Is there a trend in unemployment rate in Barcelona?

The total number of Unemployment in 2020 was increased by **220%** compared to the previous year

There was a significant jump in unemployment rate

◉ **220% Increase in 2020**

```
Monthly average registered unemplyoment people in 2019 87823.91666666667
Monthly average registered unemplyoment people in 2020 193386.91666666666
Monthly average registered unemplyoment people in 2021 251231.5
```

◉ **Increased Monthly Average Unemployment Rate**

(2019->2020) : 220%

(2020->2021) : 130%

Possible Cause can be 'Covid 19'

2 Issue

Which district has the highest unemployment rate ?

Distribution of the number of registered & who applied for the registration per each district
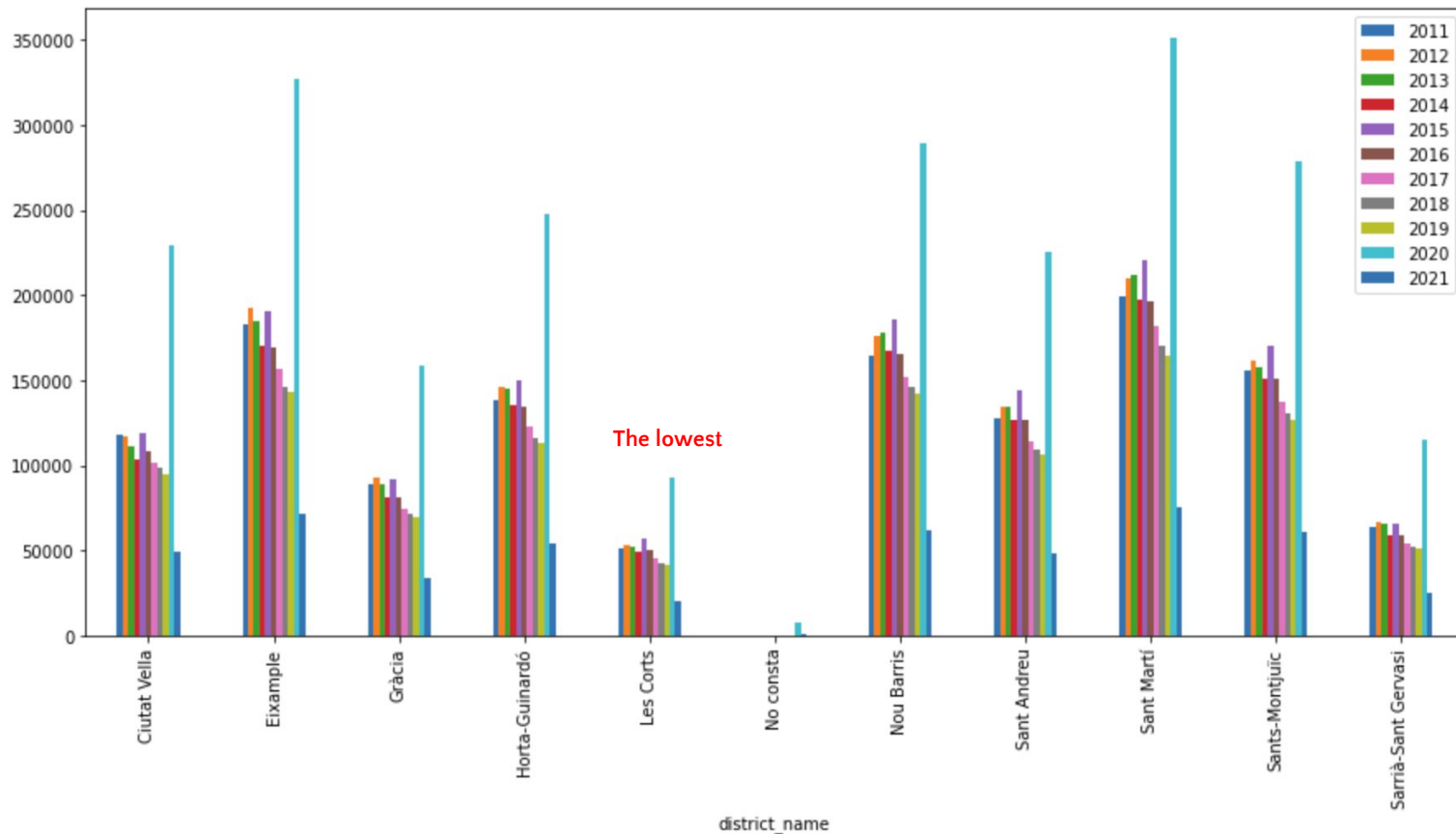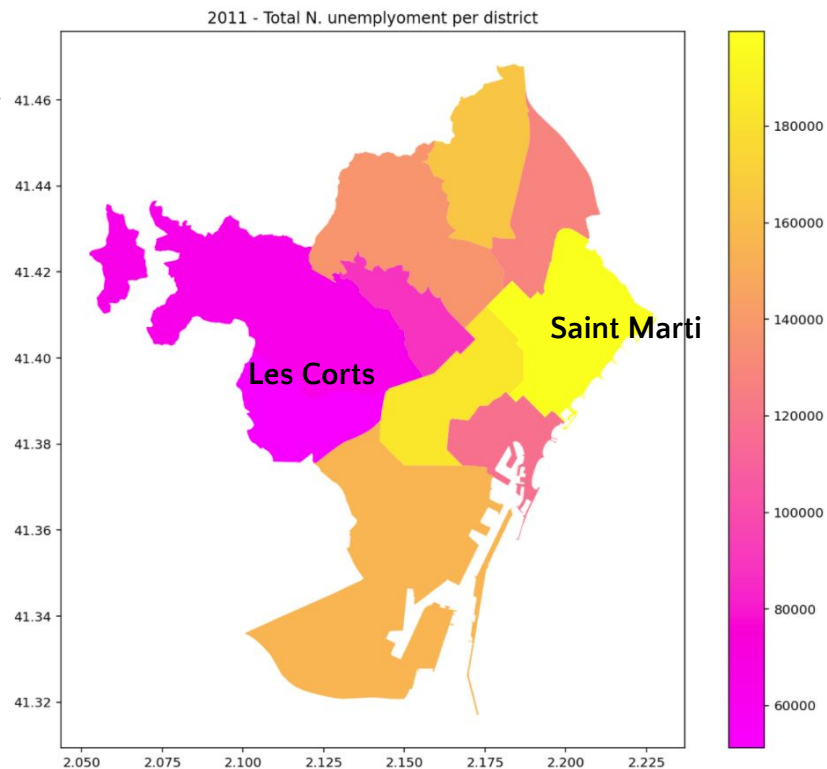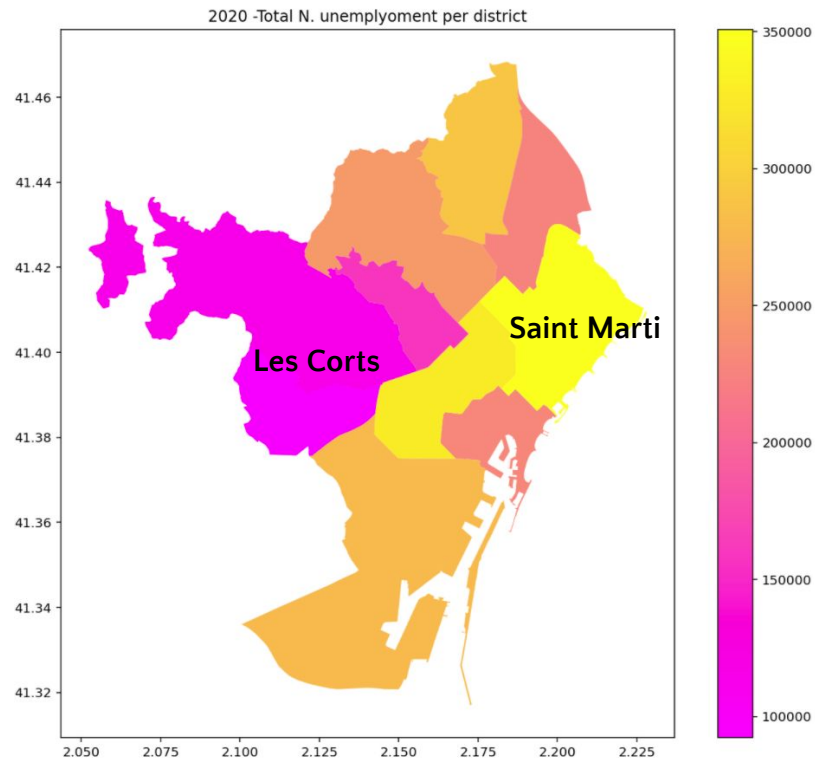
Lowest/highest number(s) of unemployment

- ◎ 'Les Corts' has the lowest number of unemployed people between 2011–2021

- ◎ 'Sant Marti' has the highest number of unemployed people between 2011–2121

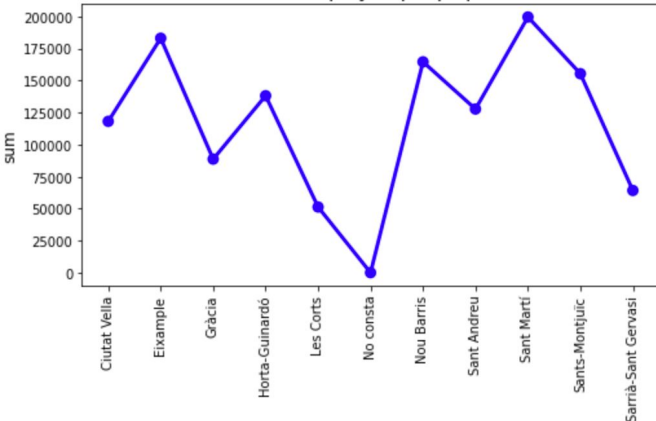# Lowest/highest number(s) of unemployment per district
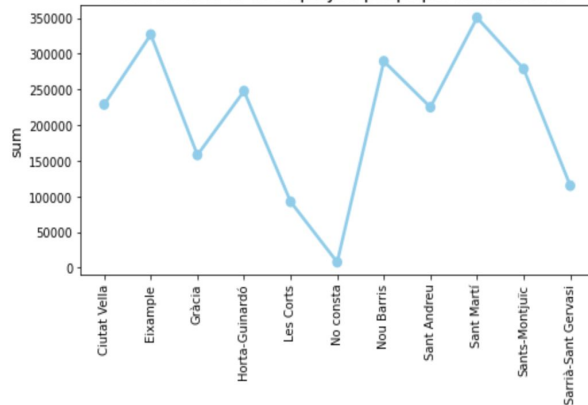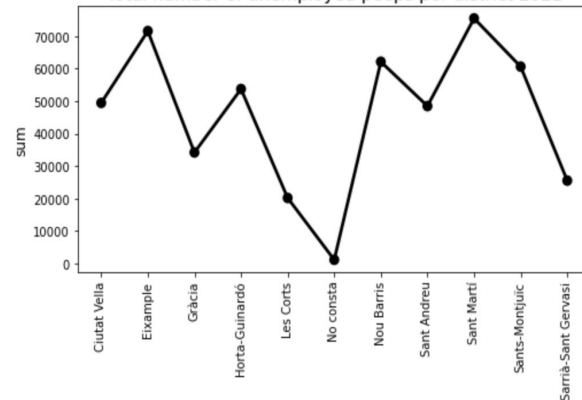

2011:

2020:

Trends

Total number of unemployed peeps per district 2011

Total number of unemployed peeps per district 2020
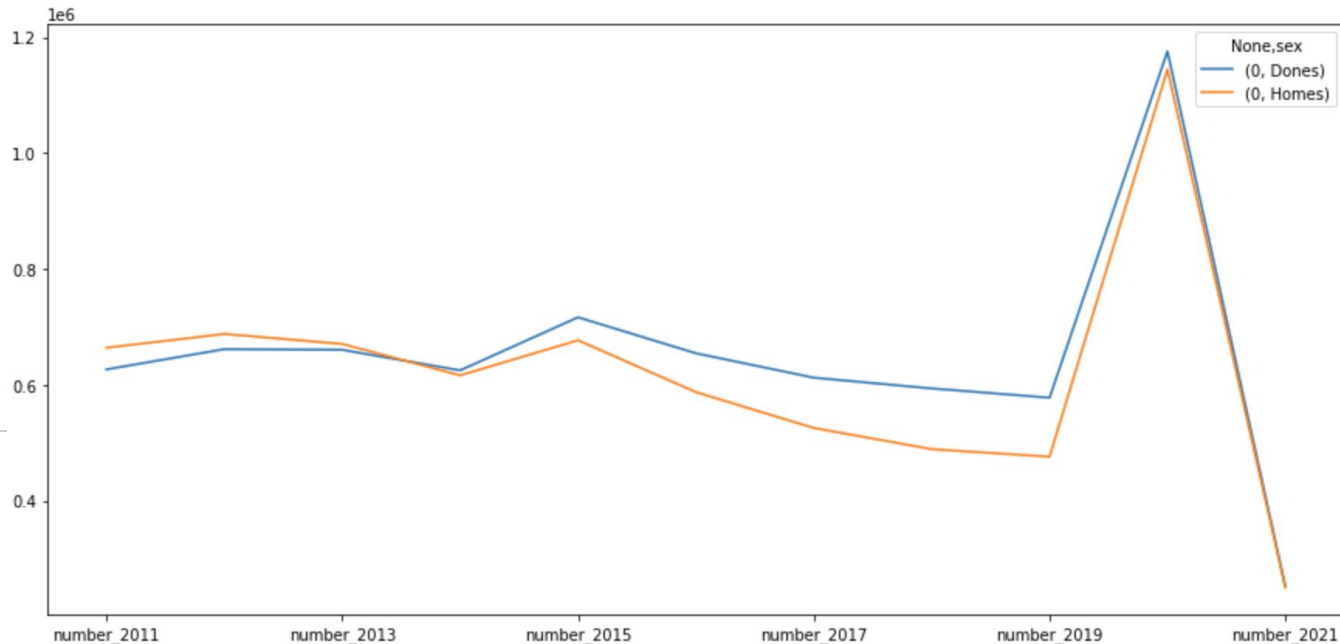

Total number of unemployed peeps per district 2021

◉ **As you can see, the unemployment trend has not been changed much last 10 years but there is a change in terms of the total number of unemployed people per district**

15

3 Issue

The difference between the total numbers of unemployment of men and women

MEN vs Women's Unemployment


```
gender_pivot_sum=gender_pivot.sum()
gender_pivot_sum=pd.DataFrame(gender_pivot_sum)
gender_pivot_sum=gender_pivot_sum.unstack()
gender_pivot_sum
```

|		0
sex	Dones	Homes
number_2011	626758	664140
number_2012	661707	687946
number_2013	660552	670661
number_2014	625114	616477
number_2015	716721	677017
number_2016	654554	587530
number_2017	612465	525644
number_2018	593870	489058
number_2019	577869	476018
number_2020	1175900	1144743
number_2021	252223	250240

◉ Except for the years [2011,2012,2013], the women tend to face more unemployment
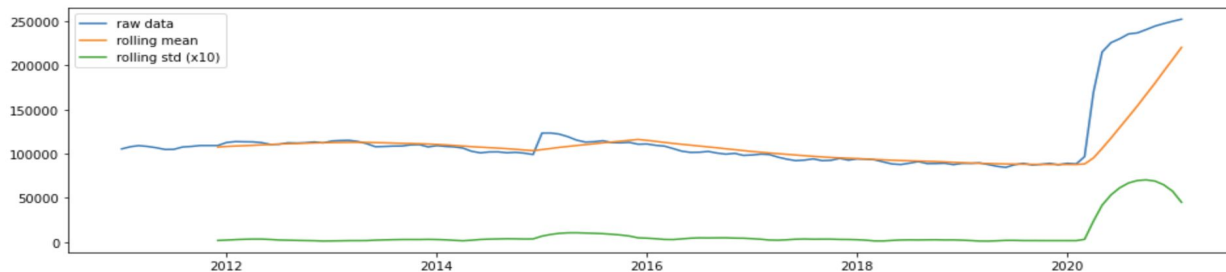
Issue

4

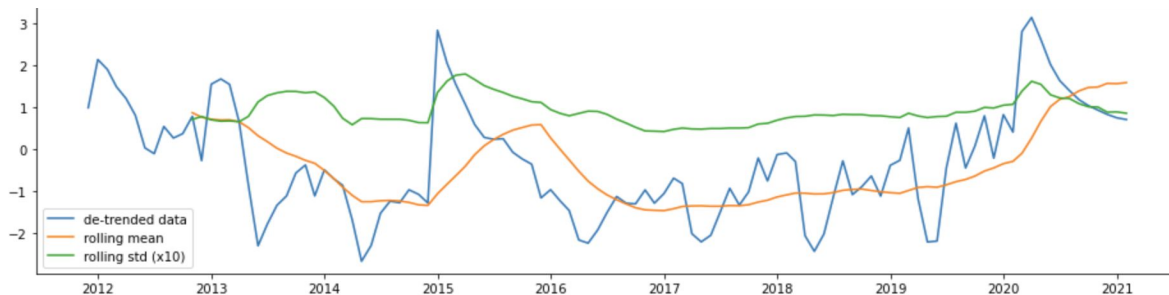Can we predict the next few month's unemployment rate?

Decompose a dataset
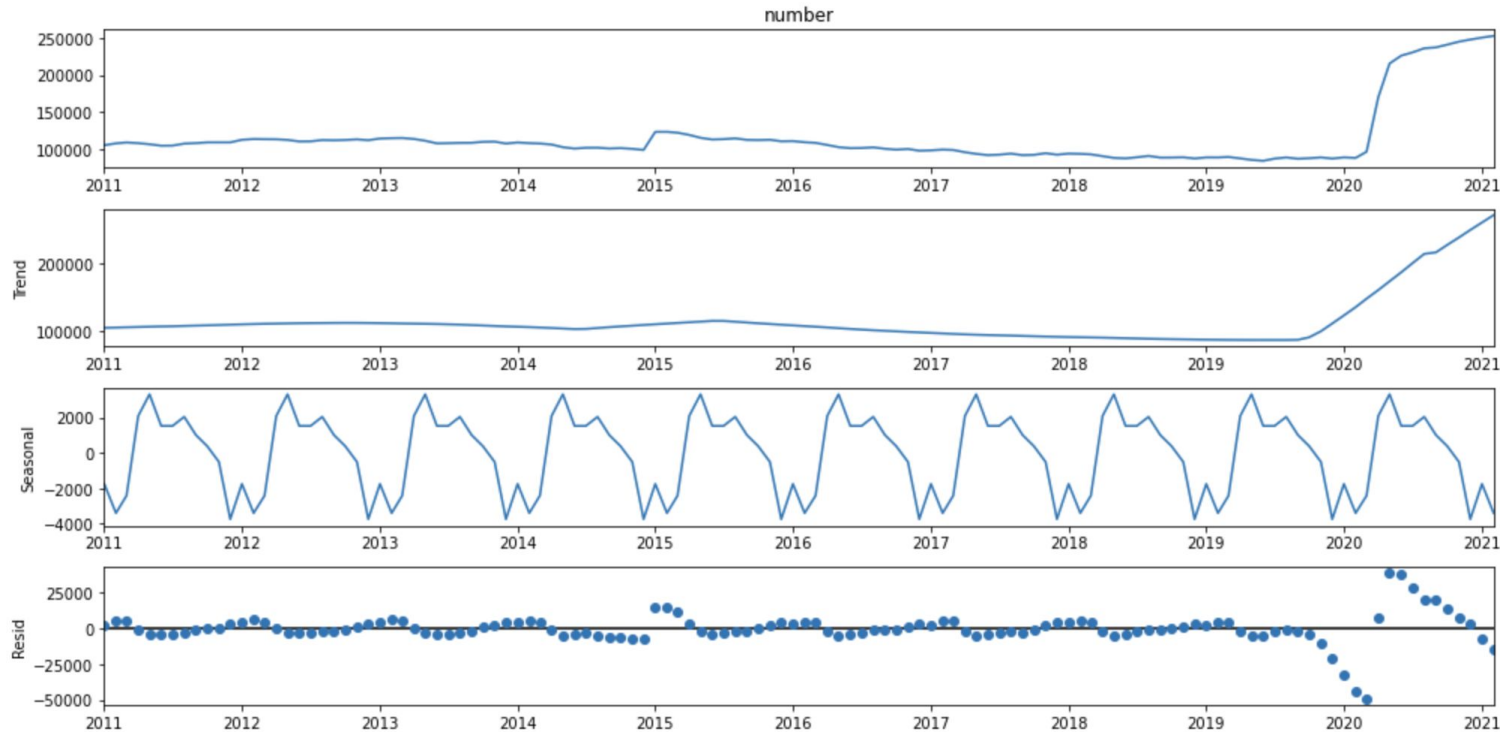
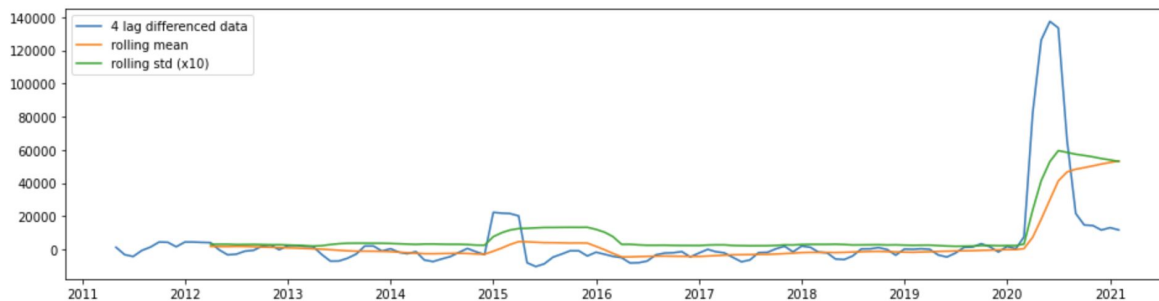- Clear trend (ups & downs) and we can see a jump in 2020

Decompose a dataset

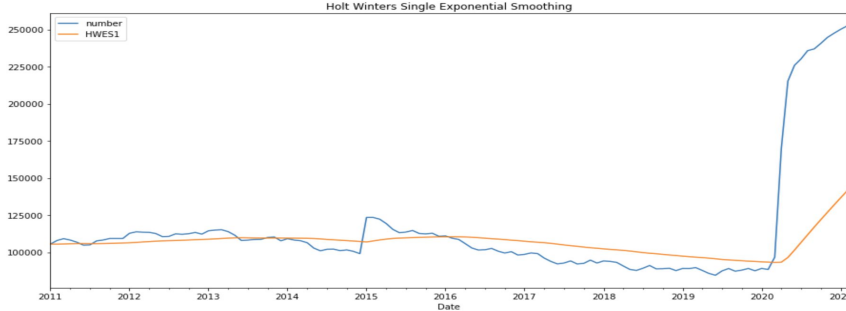

Rolling mean:

Data is non stationary

Rolling window (12):

Data is non stationary
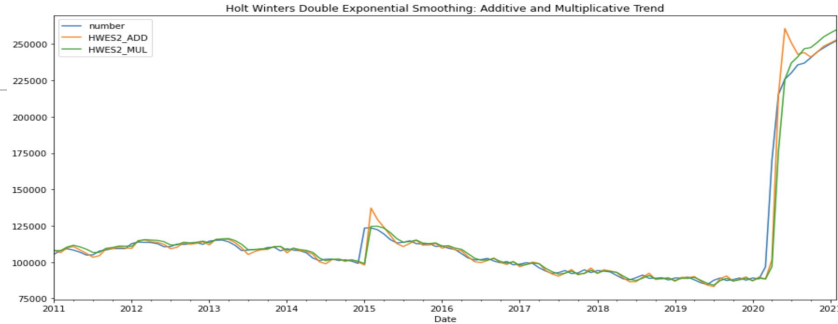
4 lag difference (12):

Data is still non stationary

Holt winters exponential smoothing

Holt Winters Single Exponential Smoothing

Holt Winters Double Exponential Smoothing: Additive and Multiplicative Trend

Triple HWES ->
(Additive and
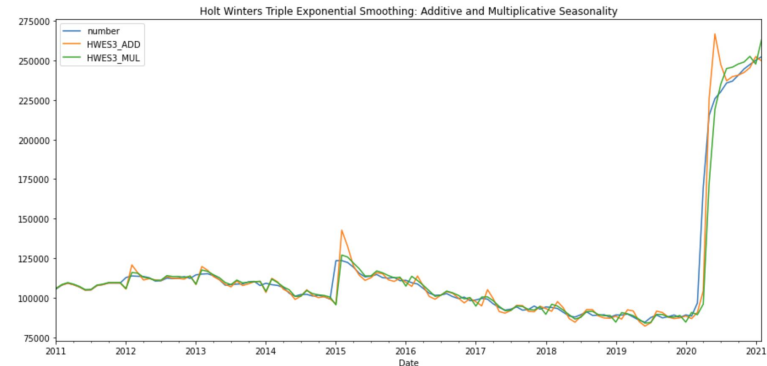 **Multiplicative Seasonality o)**

Exponential smoothing methods including Holt–Winters
 methods are appropriate for (some kinds of)
non–stationary data

<– Single HWES
(seasonality x, trends x)

<–Double HWES
(seasonality x, Additive and Multiplicative Trend o)

Holt Winters Triple Exponential Smoothing: Additive and Multiplicative Seasonality
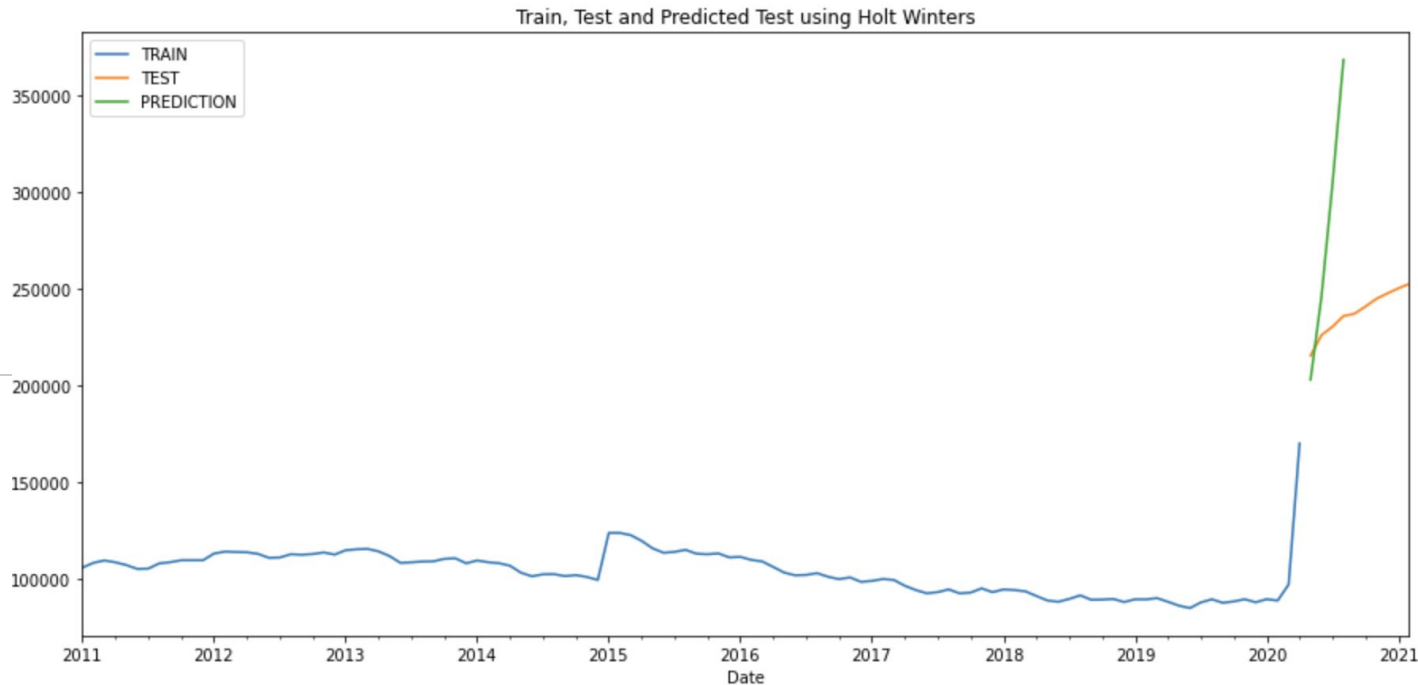
Predicting the unemployment rate

Train, Test and Predicted Test using Holt Winters

The model did not perform well as it picked the sudden jump in 2020, the prediction also shows the drastic increase in the total number of unemployment

Findings

Through EDA...

Which year?

The total number of unemployment was increased drastically in 2020.

In early 2021, the number of unemployed people was still increasing.

Which district?

The unemployment trend over 10 years hasn't been changed significantly

'Les Corts'

–> lowest number of unemployed

'Sant Marti'

–> has the highest number of unemployed people

Which Gender?

Between 2011–2013, men's unemployment rate was higher than women, since 2014 the trend has changed completely.

Total N. of Women's Unemployment

>

Total N of. Men's unemployment

Can you predict?

Sort of, yes but not really accurate model, but it seems like the unemployment number will still grow.

Next Things to do...

–Add covid data

–Add multiple features to time series model

Thanks!

Any *questions* ?

You can find me at

- https://www.linkedin.com/in/vivian-kim-89aba9199/
- viviankim621@gmail.com