

Estudio del impacto de la tasa de variación del PIB en la tasa de paro

Exploratory Data Analysis (EDA)
Por Jose Antonio Moreno De Castro
Mayo 2022



1. Objetivo

Analizar el impacto que tiene la tasa de variación del PIB sobre la tasa de paro en España tanto a nivel general como en diferentes grupos de población:

- Menores de 25 años (Tasa de paro juvenil).
- Mayores de 25 años.
- Hombres
- Mujeres

1.1 Preguntas a responder

- Relación entre la tasa de variación del PIB y la tasa de paro
- Grupo mas afectado por el paro
- Grupo menos afectado por el paro
- Grupo mas sensible a las variaciones del PIB (menor estabilidad laboral)
- Grupo menos sensible a las variaciones del PIB (mayor estabilidad laboral)

2. Recursos utilizados

- **Lenguaje de Programación: Python 3.9.5**
- **Librerías: Numpy ,Pandas y Matploblib.**
- **Jupyter Notebooks con Visual Studio Code:**
- **Microsoft Word:**
- **Adobe Acrobat Reader DC:**
- **Microsoft Power Point:**

3. BBDD

Dos bases de datos, ambas sacadas de la página oficial del Instituto Nacional de Estadística (INE)
<https://inecovid.inecovid.es/>

1. Tasas de Paro: datos de las tasas de paro en España de los últimos 20 años divididas en trimestres
2. Variación trimestral del PIB: datos de la tasa de variación trimestral del PIB de los últimos 20 años

4. Análisis

- **Exploración y descripción del Database**
- **Limpieza y preparación de datos**
- **Análisis y visualización gráfica**

4.1 Exploración y descripción del Database

- Observo los dataset para echar un vistazo a su contenido y poder detectar posibles errores que me puedan causar problemas futuros.

	Periodo	Comunidad	Total	Menores_25	Mayores_25	Parados_hombres	Parados_mujeres
0	2021T4	Total_Nacional	13.33	30.70	12.15	11.79	15.04
1	2021T4	Andalucía	20.18	39.15	18.87	17.21	23.67
2	2021T4	Aragón	9.02	19.61	8.28	8.17	9.96
3	2021T4	Asturias	10.01	10.43	10.00	9.67	10.39
4	2021T4	Islas_Baleares	14.90	32.04	13.95	14.19	15.68
...
1595	2002T1	Navarra	6.03	13.51	5.03	3.67	9.68
1596	2002T1	País_Vasco	9.56	18.10	8.64	6.20	14.39
1597	2002T1	La_Rioja	7.05	15.05	6.04	5.75	9.41
1598	2002T1	Ceuta	7.01	23.28	4.80	4.57	12.12
1599	2002T1	Melilla	10.30	26.66	7.55	6.21	18.12

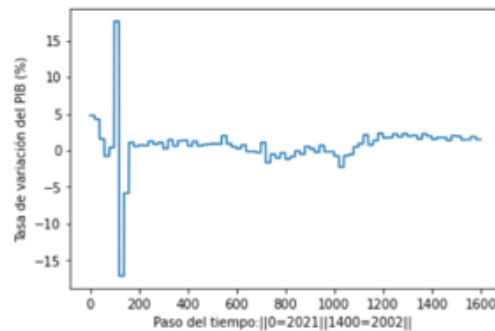
1600 rows × 7 columns

	Periodo	Var_PIB
0	2021T4	4.7583
1	2021T4	4.7583
2	2021T4	4.7583
3	2021T4	4.7583
4	2021T4	4.7583
...
1595	2002T1	1.5069
1596	2002T1	1.5069
1597	2002T1	1.5069
1598	2002T1	1.5069
1599	2002T1	1.5069

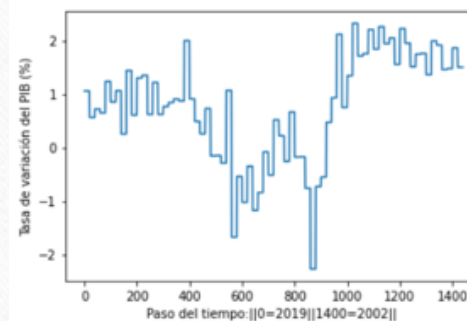
1600 rows × 2 columns

4.2 Limpieza y preparación de datos

- Juntar ambos dataset
- Eliminar la columna Periodo y Unnamed: 0
- Dividir la columna Periodo en año y trimestre
- Eliminar los datos del año 2020 y 2021 ya que eran atípicos debido al COVID



Con 2021 y 2020



Sin 2021 y 2020

4.3 Análisis y visualización gráfica

- Una vez tengo ya listo mi dataset final me pongo a analizar los datos de los que dispongo para que me den una idea inicial de lo que me va a salir como resultado:

	Total	Menores_25	Mayores_25	Parados_hombres	Parados_mujeres	Var_PIB
count	1440.000000	1440.000000	1440.000000	1440.000000	1440.000000	1440.000000
mean	16.050646	35.158062	14.299319	14.219944	18.641660	0.785196
std	7.518449	15.271719	6.993594	7.524689	8.192455	1.015171
min	0.410000	0.000000	0.460000	0.000000	1.210000	-2.267600
25%	10.050000	22.212500	8.637500	7.907500	12.537500	0.150125
50%	14.795000	33.675000	13.230000	12.980000	16.810000	0.867850
75%	21.005000	46.625000	18.975000	19.560000	23.715000	1.529675
max	39.190000	90.350000	37.210000	39.890000	55.280000	2.330400

4.3 Análisis y visualización gráfica

Primeras impresiones:

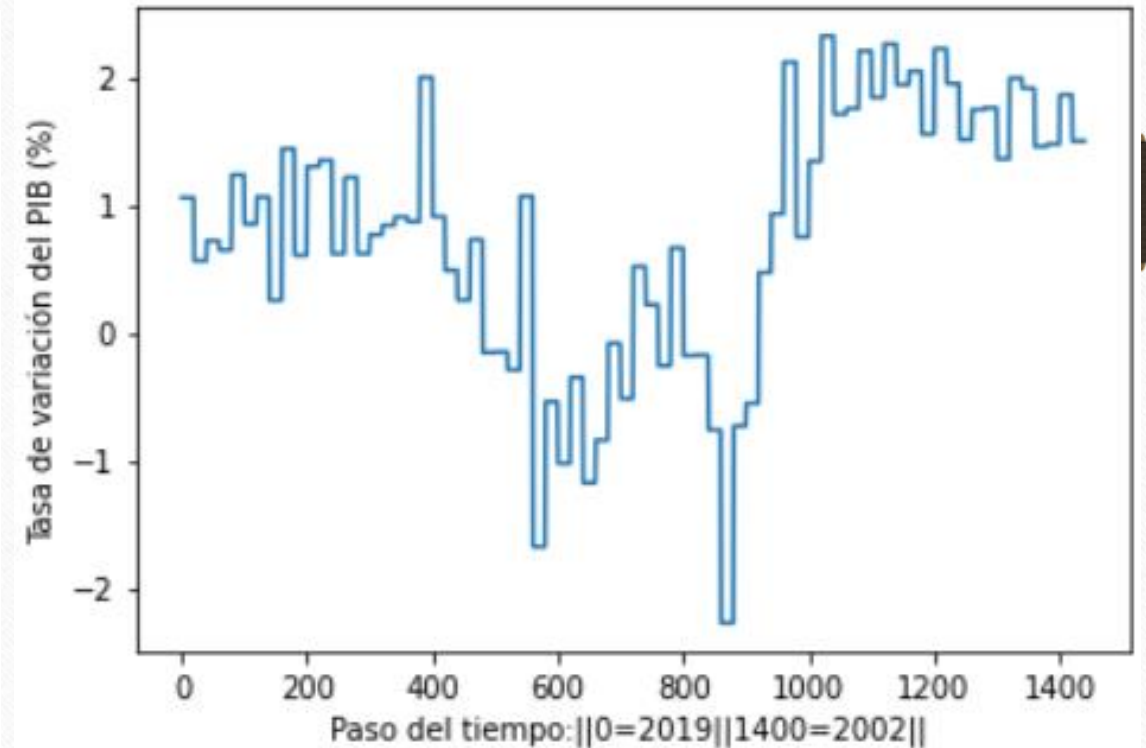
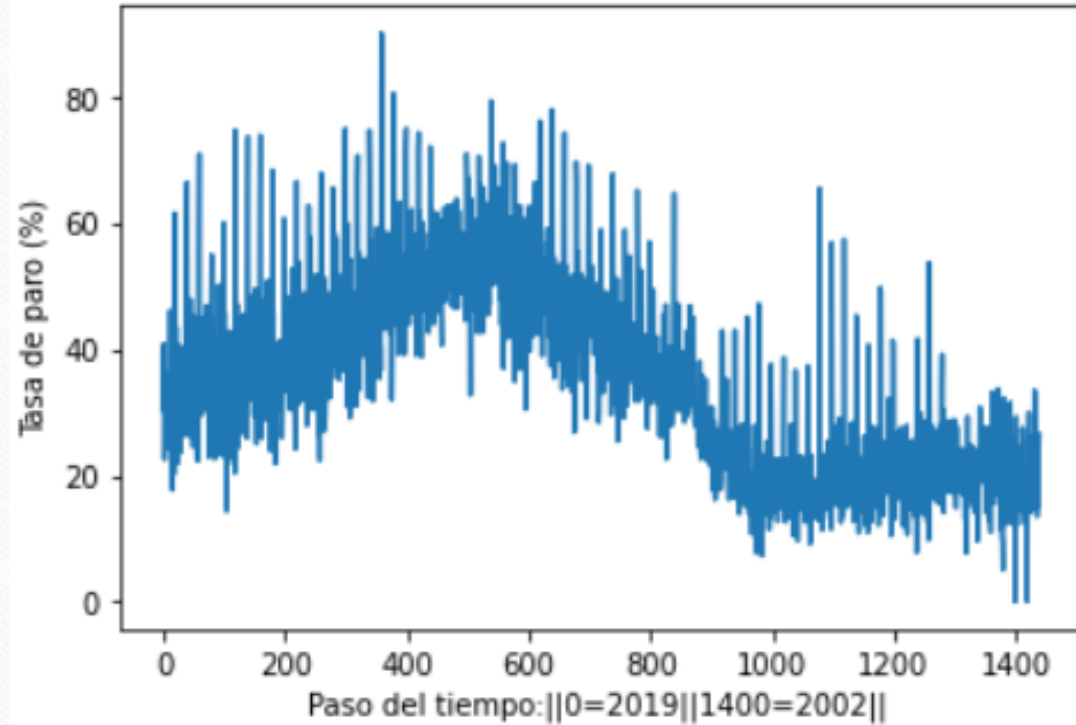
- El grupo poblacional que mas sufre el desempleo son los menores de 25 años
- El grupo poblacional de los hombres es el que menos sufre el desempleo por su tasa de paro media

	Total	Menores_25	Mayores_25	Parados_hombres	Parados_mujeres	Var_PIB
count	1440.000000	1440.000000	1440.000000	1440.000000	1440.000000	1440.000000
mean	16.050646	35.158062	14.299319	14.219944	18.641660	0.785196
std	7.518449	15.271719	6.993594	7.524689	8.192455	1.015171
min	0.410000	0.000000	0.460000	0.000000	1.210000	-2.267600
25%	10.050000	22.212500	8.637500	7.907500	12.537500	0.150125
50%	14.795000	33.675000	13.230000	12.980000	16.810000	0.867850
75%	21.005000	46.625000	18.975000	19.560000	23.715000	1.529675
max	39.190000	90.350000	37.210000	39.890000	55.280000	2.330400

4.3 Análisis y visualización gráfica

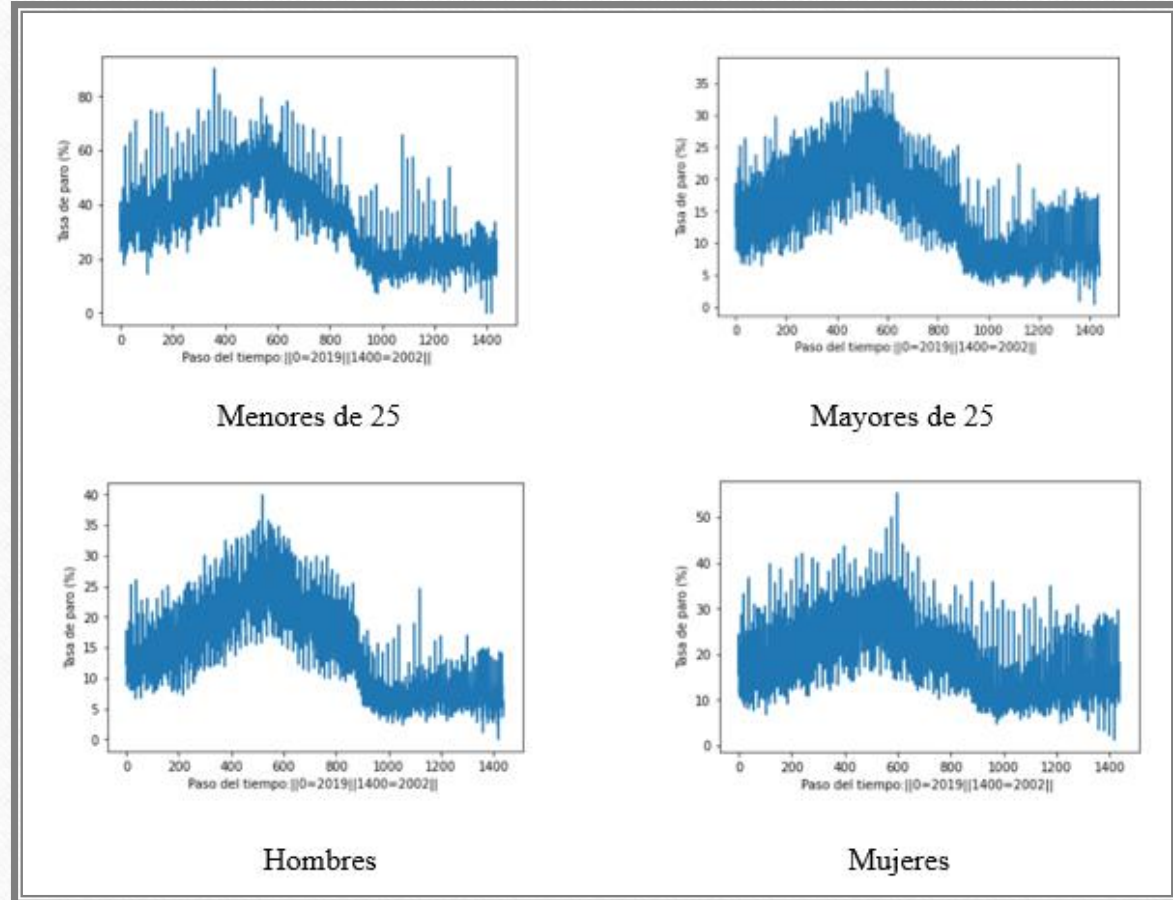
- Para continuar, calculo la covarianza y el coeficiente de correlación entre el paro total y la tasa de variación del PIB, la cual, como era de esperar, es negativa, lo que quiere decir que están relacionadas de manera inversa, y de hecho así puede verse en los gráficos, ya que el gráfico de la tasa de paro es cóncavo y el de la tasa de variación es convexo, cuando una variable sube, la otra baja:

4.3 Análisis y visualización gráfica



4.3 Análisis y visualización gráfica

- A parte de ese gráfico se sacaron los gráficos de todos los grupos poblacionales, y todos tenían la misma forma:

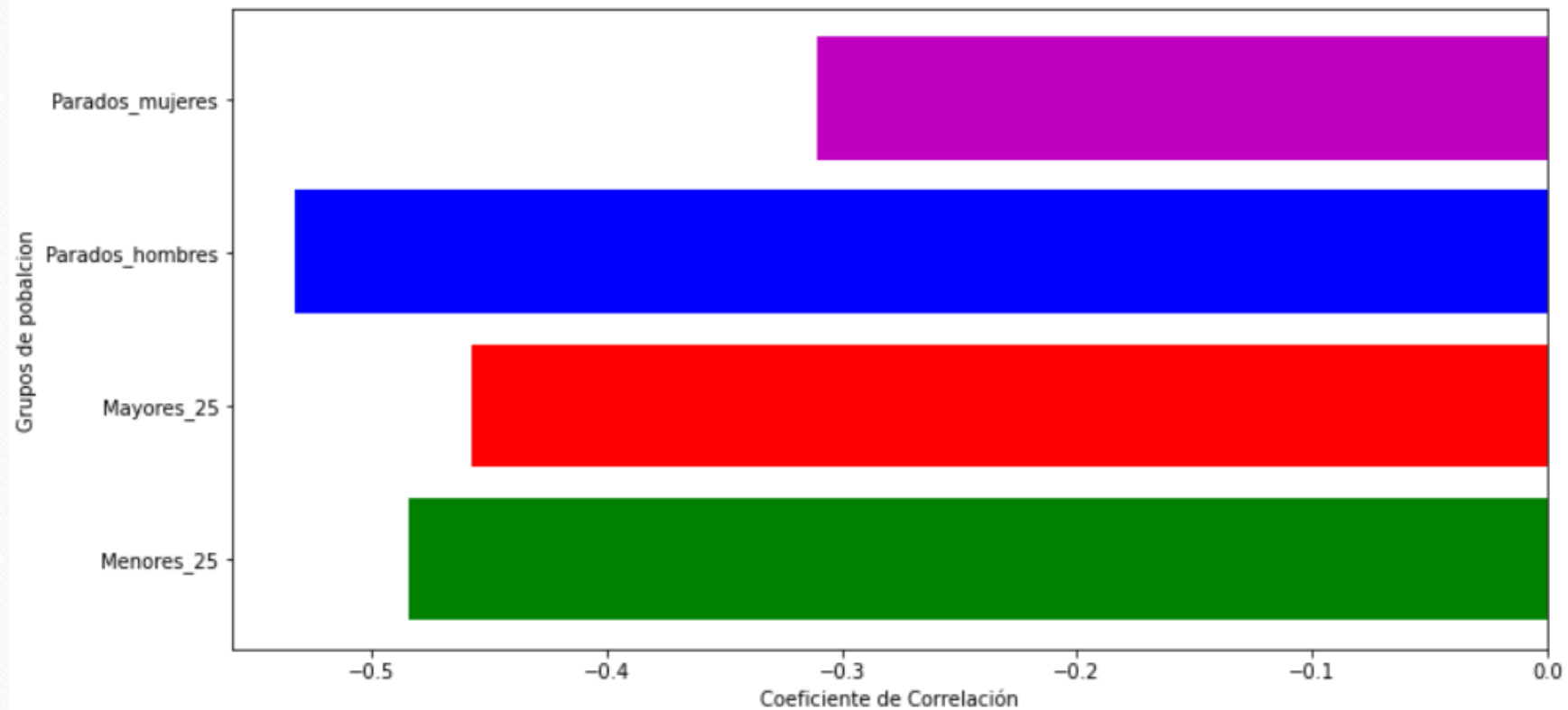


4.3 Análisis y visualización gráfica

Para continuar el análisis lo que hice fue calcular los coeficientes de correlación de cada uno de los grupos entre la variación del PIB y la tasa de paro:

- Para el grupo Menores_25 el coeficiente de correlación es de -0.4844075817740141
- Para el grupo Mayores_25 el coeficiente de correlación es de -0.45786141247544593
- Para el grupo Parados_hombres el coeficiente de correlación es de -0.5328564800617882
- Para el grupo Parados_mujeres el coeficiente de correlación es de -0.3108996031365061

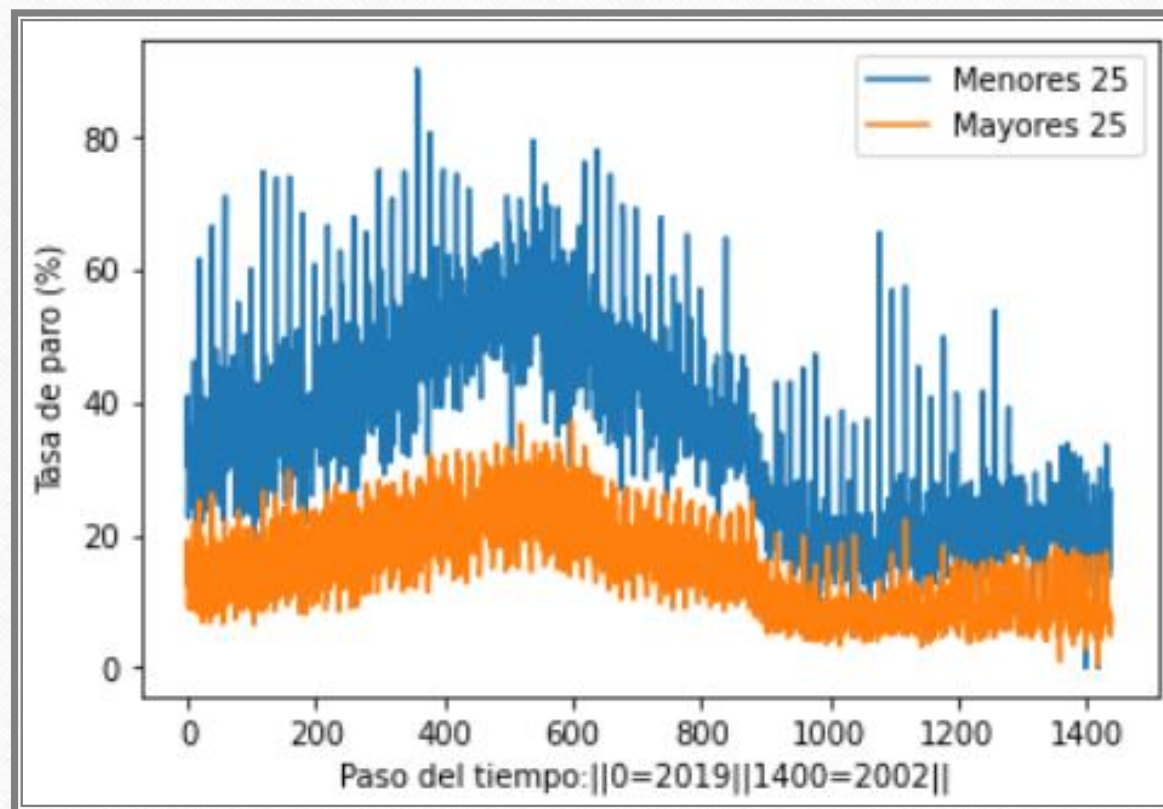
4.3 Análisis y visualización gráfica



4.3 Análisis y visualización gráfica

Comparación mayores y menores de 25:

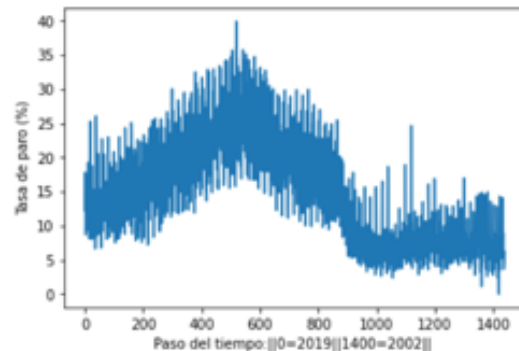
- Ante una bajada de la tasa de variación del PIB, la tasa de paro de los menores de 25 aumentará más que la de los mayores de 25.



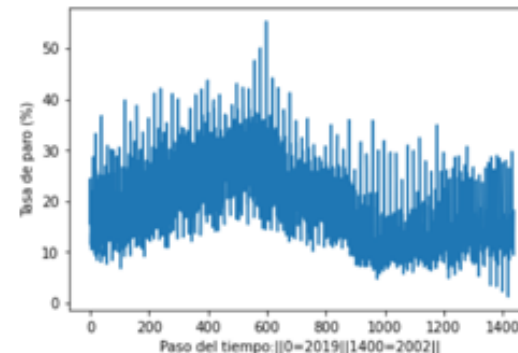
4.3 Análisis y visualización gráfica

Comparación por sexos:

- Ante una bajada de la tasa de variación del PIB, la tasa de paro de los hombres aumentará más que la de las mujeres.



Hombres



Mujeres

5. Conclusiones

- **5.1 Grupo más afectado por el paro: Menores de 25**
- **5.2 Grupo menos afectado por el paro: Mayores de 25**
- **5.3 Grupo más sensible a las variaciones del PIB: Hombres**
- **5.4 Grupo menos sensible a las variaciones del PIB: Mujeres**

Comparación con las impresiones iniciales:

- El grupo mas afectado por el paro son los menores de 25 ✓
- El conjunto poblacional masculino es el mejor de todos laboralmente ✗