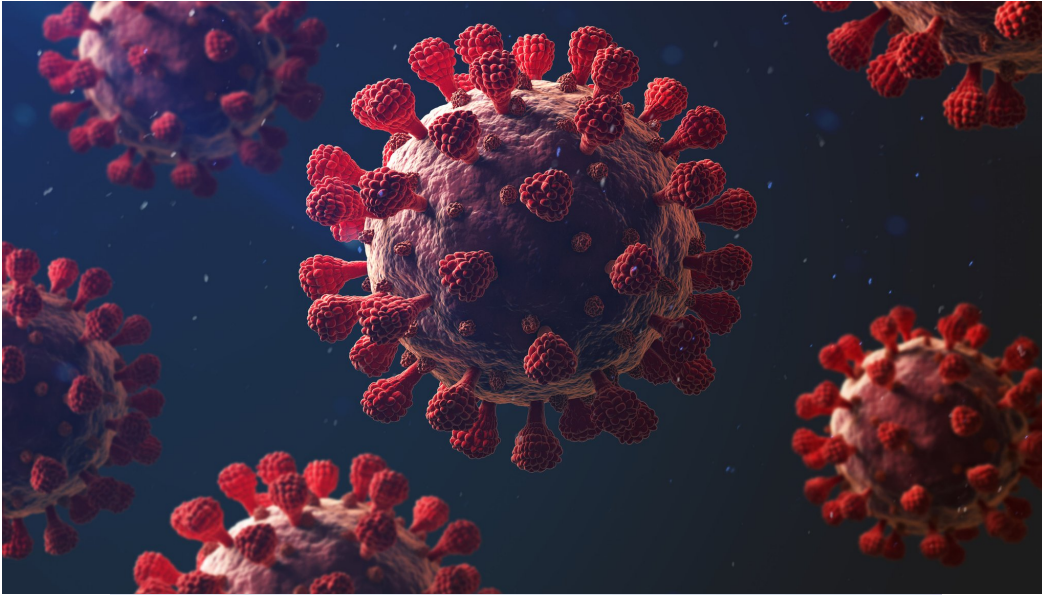


ESTUDO DA RELACIÓN ENTRE OS CASOS DE COVID-19 E A CONTAMINACIÓN DO AIRE NAS CIDADES DE VIGO E VALLADOLID

PROXECTO DE ESTATÍSTICA IES ROU 2022



**Iván Iglesias Campos, Sara Vicente Lorenzo,
Nora Ramallo Martínez e Antón Justo Vicente**

FEBREIRO-MARZO-ABRIL-MAIO 2022

3º ESO C, IES ROU, MATEMÁTICAS ORIENTADAS ÁS ENSEÑANZAS ACADÉMICAS

TITORES: ADELA ROSA RODRÍGUEZ RODRÍGUEZ e DANIEL FERNÁNDEZ VENCE

ÍNDICE

1. INTRODUCCIÓN	3
2. OBXECTIVO DA INVESTIGACIÓN	3
2.1 OBXECTIVO	3
2.2 SUBOBXECTIVOS	3
3. POBOACIÓN E VARIABLES	4
3.1 POBOACIÓN	4
3.2 VARIABLES	4
4. DESCRICIÓN DA TOMA DE DATOS	4
5. ANÁLISE ESTADÍSTICA DOS DATOS	6
5.1 TÁBOAS DE FRECUENCIAS	6
5.1.1 DATOS MÁXIMOS E MÍNIMOS DAS DIFERENTES VARIABLES	6
5.1.2 EXEMPLO DE TÁBOA DE FRECUENCIAS	7
5.2 GRÁFICAS	8
5.2.1 PRIMEIRA VAGA	8
5.2.2 SEGUNDA VAGA	9
5.2.4 CUARTA VAGA	11
5.2.5 QUINTA VAGA	12
5.2.6 SEXTA VAGA	13
5.3 TÁBOAS DE ANÁLISE DAS MEDIDAS DE CENTRALIZACIÓN E DISPERSIÓN	14
5.3.1 CASOS POSITIVOS DE COVID-19	14
5.3.2 CONTAMINACIÓN DO AIRE EN PM25	16
5.4 ESTUDO DE CORRELACIÓN	18
5.4.2 SEGUNDA VAGA	19
5.4.3 TERCEIRA VAGA	20
5.4.4 CUARTA VAGA	21
5.4.5 QUINTA VAGA	22
6. CONCLUSIÓN	24
7. PROPOSTAS DE MELLORA	24
8. REFERENCIAS	24

1. INTRODUCCIÓN

Hai dous anos un novo virus introduciuse nas nosas vidas e cambiaríanolas por completo pechándonos nas nosas casas durante case tres meses e introducindo a máscara no noso día a día, sabemos que estamos falando do SARS-CoV-2, que procova a enfermidade coñecida como COVID-19.

Aínda que levemos xa dous anos convivindo con este virus que se detectou na China en decembro do 2019 descoñecemos máis do que sabemos sobre el, por iso, a día de hoxe, séguense a facer investigacións científicas relacionadas con este microorganismo.

Moitas destas investigacións teñen que ver coa forma de transmisión do virus e cos parámetros que teñen influencia nesa propagación, como pode ser a contaminación no aire. A influencia da contaminación no aire e a temperatura atmosférica na transmisión do patóxeno¹ foi o obxecto de estudo de tres investigadores: Elena Salamanca Fernández, Miguel Rodríguez Barranco e María José Sánchez, pertencentes os tres á Escola Andaluza de Saúde Pública (EASP) (Granada), ao Instituto de Investigación Biosanitaria (ibs.GRANADA) (Granada) e ao Centro de Investigación Biomédica en Rede de Epidemioloxía e Saúde Pública (CIBERESP) (Madrid). A parte de pertencer a estas tres entidades María José Sánchez tamén forma parte do Departamento de Medicina Preventiva e Saúde Pública da Universidade de Granada (Granada).

Os tres investigadores tomaron os datos de poboacións en China, nos Estados Unidos e en Europa, concretamente en Italia. Nós, dentro das nosas posibilidades, que son moito máis limitadas cás destes investigadores, levamos esta investigación ao noso entorno, á cidade de Vigo, pero tiñamos que comparala con outra cidade, neste caso de interior, xa que Vigo é unha cidade costeira. A nosa elección foi a da cidade de Valladolid, pois ten un número de habitantes moi similar a Vigo (ao redor dos 300.000).

Deste xeito, e baseándonos no estudo destes investigadores, realizamos un traballo de investigación no que tentaremos establecer relacións entre os casos positivos en COVID-19 durante a pandemia e a contaminación do aire en partículas de PM25 (unha das medidas de contaminación que utilizan os investigadores no seu estudo) durante a mesma, e que será presentado a parte de ao XII Concurso Incubadora de Sondaxes e Experimentos da Sgapeio ao II congreso de Estatística do IES ROU.

2. OBXECTIVO DA INVESTIGACIÓN

2.1 OBXECTIVO

Establecer unha relación entre a contaminación do aire cos casos diagnosticados de COVID-19 nas áreas de influencia sanitaria da cidade costeira de Vigo e na cidade interior de Valladolid e concluír se os parámetros están relacionados estudando se un afecta ó outro.

2.2 SUBOBXECTIVOS

1. Estudar os datos de casos positivos en COVID-19 das áreas de influencia sanitaria das cidades de Vigo e Valladolid.
2. Estudar os datos de contaminación en PM25 das dúas áreas de influencia sanitaria.
3. Realizar táboas de frecuencia e gráficos.
4. Estudar as medidas de centralización e de dispersión.
5. Establecer as relacións entre ámbalas variables.

3. POBOACIÓN E VARIABLES

3.1 POBOACIÓN

A nosa poboación descríbese como o número de días nos que se detectan casos positivos de COVID-19 e contaminación no aire en partículas de PM25.

3.2 VARIABLES

Variables	Cuantitativas discretas	Cuantitativas continuas	Cualitativas
Número de casos de COVID-19 nas diferentes áreas sanitarias (Vigo e Valladolid)			
Contaminación en partículas de menos de 2,5 micras no aire en cada día das diferentes vagas pandémicas nas dúas cidades (Vigo e Valladolid)			

4. DESCRICIÓN DA TOMA DE DATOS

Xa que o noso obxectivo é o de estudar a relación entre os casos de COVID-19 e a contaminación do aire interézanos comparar os datos de COVID-19 da área sanitaria de Vigo cos datos tamén de COVID-19 de Valladolid, cuxa área de influencia é a propia provincia de Valladolid, segundo a contaminación do aire en ambas cidades. Estes datos foron obtidos de bases de datos de carácter público. Escollemos estas dúas áreas de influencia xa que a cidade de Vigo é costeira e a cidade de Valladolid é unha cidade de interior, e ambas teñen un número semellante de habitantes, ao redor dos 300.000 (como se explica na introdución).

Para o noso estudo a nosa hipótese é que existirá unha relación entre os casos de COVID-19 de ambas áreas de influencia sanitaria coa súa contaminación en partículas de PM25, tal e como suxire a bibliografía ao efecto.

Para saber de que datas obter tanto os datos de casos de coronavirus e a contaminación obtivemos das seguintes páxinas as datas de principio a fin das vagas de COVID-19 para poder dividir o noso estudo. Esas páxinas son as seguintes:

<https://www.rtve.es/noticias/coronavirus-graficos-mapas-datos-covid-19-espana-mundo/>

<https://www.elindependiente.com/vida-sana/salud/2021/03/12/el-ano-del-covid-cronologia-de-la-pandemia-en-espana/>

O criterio para a división da pandemia en ondas baseouse en observar as gráficas presentes nas anteriores páxinas e buscar as datas de comezo de subida dos casos e da súa estabilización, deste xeito a data de subida e a data de inicio da vaga e a de estabilización e a do fin desta.

Segundo o criterio anterior a pandemia quedou dividida nas seguintes datas:

- 1ª vaga: 13/03/2020 - 31/05/2020

- 2ª vaga: 04/08/2020 - 09/12/2020
- 3ª vaga: 10/12/2020 - 12/03/2021
- 4ª vaga: 22/03/2021 - 26/05/2021
- 5ª vaga: 27/06/2021 - 16/09/2021
- 6ª vaga: 12/11/2021 - 28/02/2022

Os datos galegos correspondentes aos casos de coronavirus da área sanitaria de Vigo obtivémolos da páxina web do Sergas: <https://coronavirus.sergas.gal/datos/#/gl-ES/galicia>.

No caso dos datos dos casos da cidade de Valladolid obtivémolos da páxina da Xunta de Castela e León: <https:// analisis.datosabiertos.jcyl.es/explore/dataset/situacion-epidemiologica-coronavirus-en-castilla-y-leon/export/?disjunctive.provincia&refine.provincia=Valladolid>.

Unha vez obtidos estes datos, obtivemos os datos de contaminación en partículas de PM25, que se miden en microgramos por metro cúbico ($\mu\text{g}/\text{m}^3$). No caso de Vigo a través da páxina web de Meteogalicia: https://www.meteogalicia.gal/Caire/datos.action?request_locale=es

Para obter os datos eliximos a estación de medida de calidade do aire de Coia, xa que esta está situada nun barrio céntrico da cidade e ofrece os datos de contaminación de PM25. Os datos correspondentes a Valladolid obtivémolos da páxina de medio ambiente da Xunta de Castela e León: <http://servicios.jcyl.es/esco/cargarFrmDatosHistoricos.action>

Os datos de Castela e León correspóndese coa estación de medida da calidade do aire chamada a Rubia II, situada no barrio céntrico vallisoletán da Rubia e que mide partículas de PM25, razóns polas que eliximos esa estación.

É importante destacar que algúns datos de contaminación son nulos por mantemento ou fallos na estación de medida de calidade do aire.

Durante a recolla dos datos tivemos diferentes tipos de problemas. O primeiro problema que nos xurdiu consistiu en que os datos de casos positivos de COVID-19, no caso de Galicia e Castela e León, non se dan por cidades (que sería do noso interese) senón que se dan por áreas sanitarias (no caso de Galicia, que non se corresponden coa provincia) e por áreas de influencia (no caso de Castela e León e que se corresponde coa provincia), mentres que os datos de contaminación danse por cidades. Por iso tivemos que asumir que os datos de COVID-19 recollidos correspondíanse, na súa maioría, cos casos positivos da cidade de influencia.

O noso segundo problema viu dado polo período que deberíamos elixir para obter os datos de contaminación que podía ser ben horario ou diario. Debido a que os datos de COVID-19 son medidos de forma diaria, para aforrar traballo á hora de pasar os datos horarios a diarios, decidimos descargar os datos de tipo diario que se corresponden co dato de contaminación das 00:00 en ambas comunidades.

Unha vez resoltos estes problemas xa puidemos pasar os nosos datos a unha folla de Cálculo de Google que foi dividida en catro columnas. Na primeira colocamos as datas de rexistro de casos positivos e contaminación do aire (correspondente coa nosa poboación) e nas seguintes catro columnas colocamos os datos positivos e a contaminación do aire en partículas de PM25 en Vigo e Valladolid (correspondentes coas nosas variables).

Enlace á folla de Cálculo de Google: [Táboas de frecuencias e gráficas.xlsx](#)

5. ANÁLISE ESTATÍSTICA DOS DATOS

Unha vez recollidos todos os datos na folia de cálculo, para realizarmos a análise estatística creamos unha segunda folia de cálculo que foi dividida en seis páxinas, unha por cada vaga pandémica.

Por cada vaga tivemos que facer catro táboas de frecuencias: a primeira para os casos de COVID-19 de Vigo, a segunda para os de Valladolid, a terceira para a contaminación de Vigo e a cuarta para a de Valladolid. Deste xeito fixemos 24 táboas de frecuencia.

Tamén realizamos unha gráfica por variable en cada vaga, é dicir, nunha mesma gráfica incluimos os casos de COVID-19 das dúas cidades e en outra a contaminación en PM25 das dúas urbes. En total fixemos 12 gráficas.

Finalmente calculamos as medidas de centralización e de dispersión.

5.1 TÁBOAS DE FRECUENCIAS

Como se dixo no punto anterior realizamos unha táboa de frecuencias por cada variable en cada cidade (Vigo ou Valladolid), en total catro táboas por vaga. Para dividir os datos en intervalos, no caso dos casos de COVID-19 obtivemos o máximo número de casos e o mínimo en cada cidade e, segundo a vaga e a cidade, a amplitude dos intervalos foi de 25, 50 ou 100 casos. O mesmo fixemos coa contaminación en PM25 pero a amplitude dos intervalos foi nas dúas cidades e en todas as vagas de $5 \mu\text{g}/\text{m}^3$. (Estes datos máximos e mínimos expóñense no punto 5.1.2).

5.1.1 DATOS MÁXIMOS E MÍNIMOS DAS DIFERENTES VARIABLES

Na seguinte táboa representamos os datos máximos e mínimos das dúas variables nas dúas cidades nas seis diferentes vagas. Estes datos servíronnos para saber a amplitude de cada intervalo de datos nas diferentes cidades nas diferentes vagas.

	CASOS COVID-19				CONTAMINACIÓN PM25			
	Vigo		Valladolid		Vigo		Valladolid	
	Máximo	Mínimo	Máximo	Mínimo	Máximo	Mínimo	Máximo	Mínimo
1ª VAGA	154	0	234	0	20	0	26	0
2ª VAGA	262	3	795	0	27	0	29	0
3ª VAGA	369	16	746	11	37	0	40	0
4ª VAGA	86	12	99	11	34	0	23	1
5ª VAGA	398	9	506	4	17	1,9	28	0
6ª VAGA	2508	27	2830	24	37	0	38	0

5.1.2 EXEMPLO DE TÁBOA DE FRECUENCIAS

Nas nosas táboas de frecuencias incluimos: a Marca de Clase (X_i), que é a media do límite inferior e superior do intervalo; a frecuencia absoluta (n_i), que é o número de veces que se repiten os valores pertencentes ao intervalo (para contar este número de veces utilizamos a función `CONTAR.SI.CONJUNTO`); a frecuencia absoluta acumulada (N_i), que consiste no sumatorio de todos os valores das frecuencias absolutas menores a ese valor de frecuencia absoluta; a frecuencia relativa (f_i), que é o cocente da frecuencia absoluta entre o número total de datos; a frecuencia relativa acumulada (F_i), que é o sumatorio de todos os valores da frecuencia relativa menores a ese valor; e a porcentaxe (%), que é o produto da frecuencia relativa por 100. Por último, como medidas de centralización incluimos: a media, a moda, o intervalo modal e a mediana e como medidas de dispersión: a desviación típica e a varianza.

Outras medidas son o coeficiente de correlación (calculado só unha vez por variable en cada vaga sen depender das cidades, xa que é a relación entre os intervalos de datos de casos de COVID-19 e de contaminación das dúas cidades) e o coeficiente de variación, que é o cocente da desviación típica e a media.

Aquí expomos o exemplo dunha das táboas de frecuencia realizadas, correspondente cos casos de COVID-19 de Vigo na primeira vaga:

TÁBOA FRECUENCIA CASOS COVID-19 VIGO

INTERVALOS	Límite inferior	Límite superior	Marca de clase (X_i)	Frec absoluta (n_i)	Frec absoluta acumulada (N_i)	Frec relativa (f_i)	Frec relativa acumulada (F_i)	%	$X_i * n_i$	Centralización	Dispersión	COEFICIENTE CORRELACIÓN CASOS COVID-19	COEFICIENTE DE VARIACIÓN
[0-25)	0	25	12,5	51	51	0,64	0,64	63,75	637,5	Media	Desviación típica	0,34	1,04
[25-50)	25	50	37,5	20	71	0,25	0,89	25,00	750	27,19	28,24		
[50-75)	50	75	62,5	3	74	0,038	0,92	3,75	187,5	Moda	Varianza		
[75-100)	75	100	87,5	5	79	0,062	0,995	6,25	437,5	12,5	797,40		
[100-125)	100	125	112,5	0	79	0	0,99	0,00	0	Intervalo modal			
[125-150)	125	150	137,5	0	79	0	0,99	0,00	0	[0-25)			
[150-175)	150	175	162,5	1	80	0,012	1	1,25	162,5	Mediana			
SUMA				80		1		100	2175	12,5			

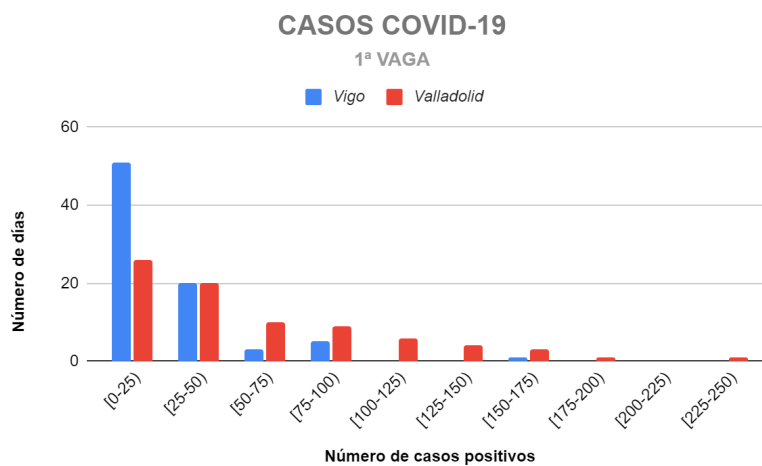
5.2 GRÁFICAS

Para realizar a análise estatística dos datos fixemos diagramas de barras no lugar de histogramas, xa que o histograma non nos permite comparar os datos das dúas cidades á vez ao non poder incluír dúas columnas nun mesmo intervalo.

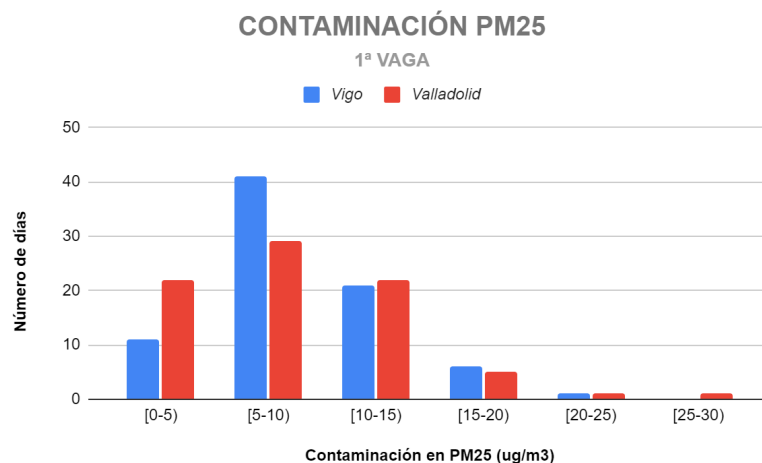
En cada gráfico no eixo X representamos os valores da variable por intervalos que poden pertencer aos casos positivos de COVID-19 ou á contaminación en PM25 dependendo da gráfica, no eixo Y representamos o número de días nos que se dan estos casos (a frecuencia absoluta).

As gráficas non están representadas en histograma xa que a folia de cálculo non permite comparar dúas variables nese formato.

5.2.1 PRIMEIRA VAGA

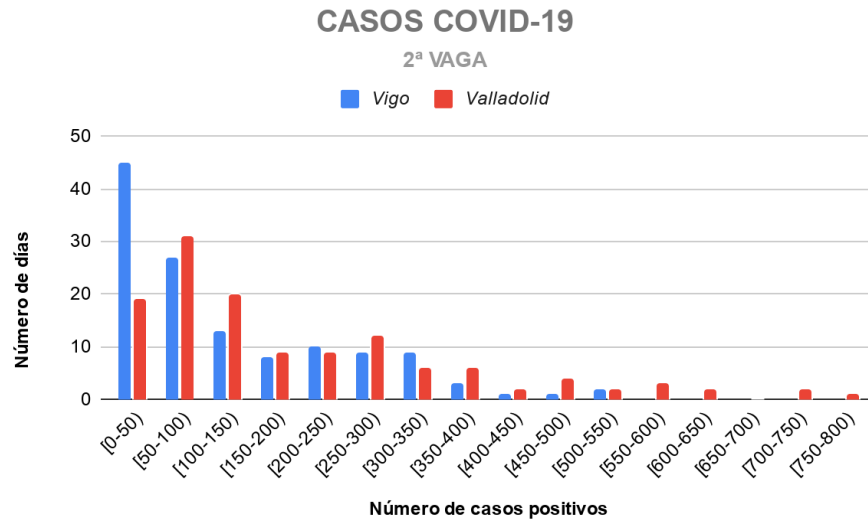


Nesta primeira gráfica están representados os casos de COVID-19 de Vigo e Valladolid durante a primeira vaga, os datos están divididos en intervalos cunha amplitude de 25 casos ata os 250, podemos observar que en Vigo hai menos casos que en Valladolid. O intervalo con máis datos é o de 0 a 25 casos, no caso de Vigo con 51 días nos que houbo entre 0 e 25 casos e no caso de Valladolid con 26 días.

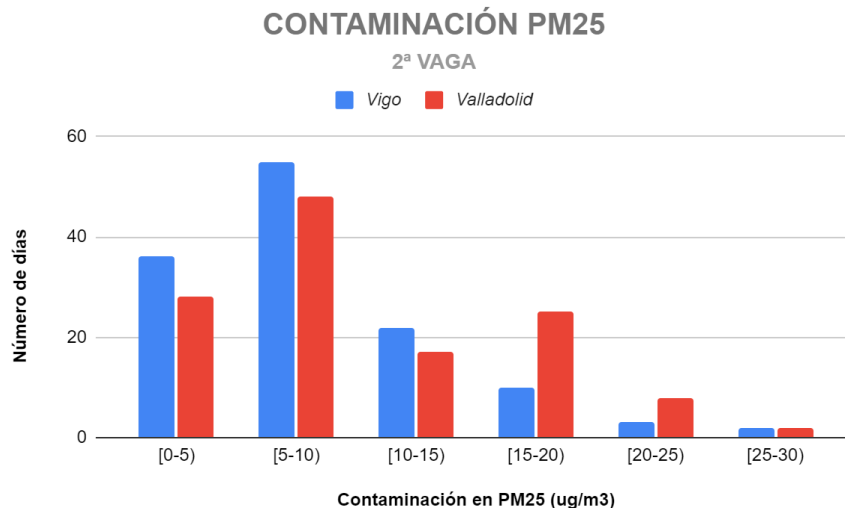


Nesta segunda gráfica representáanse os datos de contaminación en PM25 divididos en intervalos de amplitude 5, os datos chegan ata $30 \mu g/m^3$, con total de seis intervalos. O intervalo co maior número de datos é o que abrangue os datos de 5 a $10 \mu g/m^3$ con 41 días no caso Vigo e con 29 en Valladolid.

5.2.2 SEGUNDA VAGA

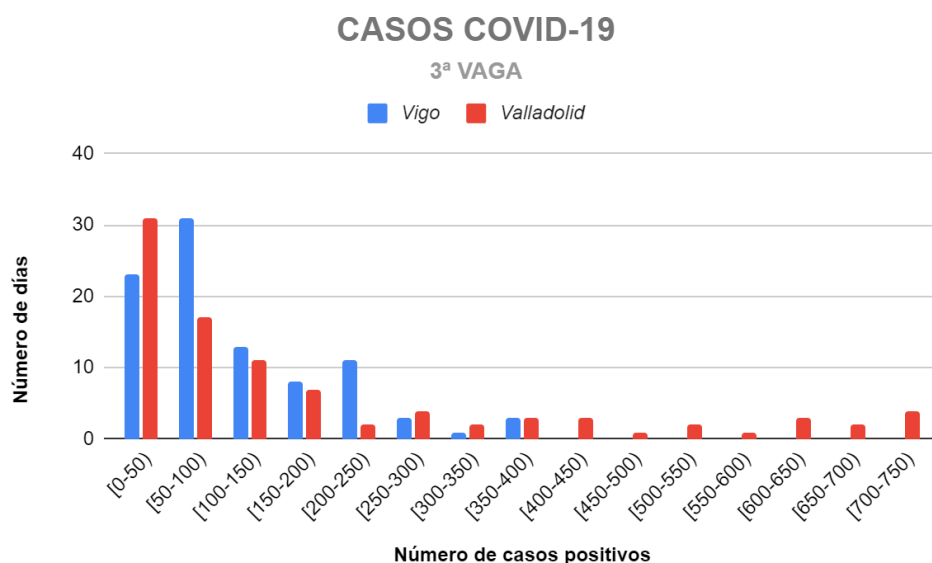


Nesta terceira gráfica están representados os casos de COVID-19 nas cidades de Vigo e Valladolid durante a segunda vaga en intervalos de amplitude de 50 casos que chegan ata os 800, nesta vaga volve a haber máis casos en Valladolid. Os intervalos con maior número de datos son diferentes nas dúas cidades, en Vigo é o de 0 a 50 casos con 45 días e en Valladolid é o de 50 a 100 casos con 50 días.

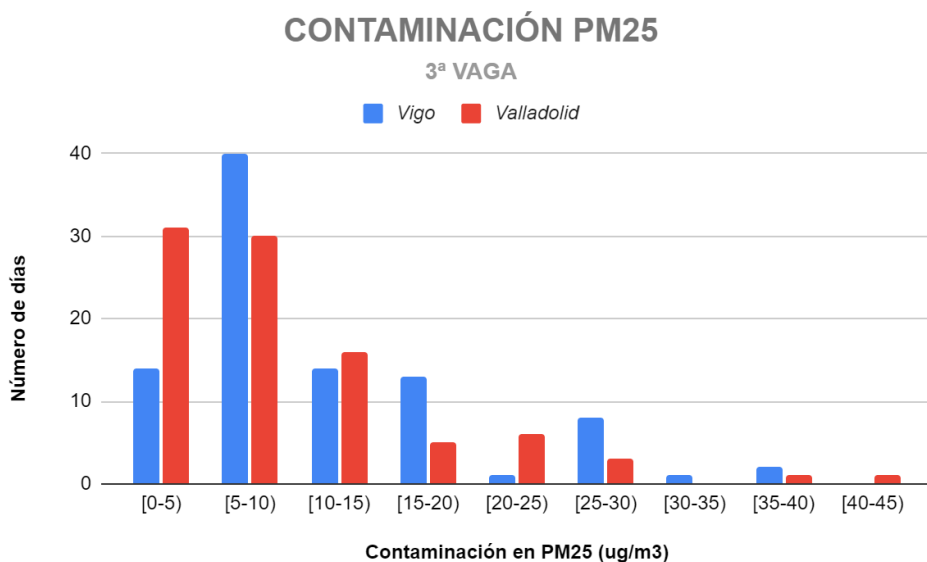


Neste cuarta gráfica están representados os casos de contaminación en Vigo e Valladolid durante a segunda vaga, divididos en intervalos de amplitude 5, ata chegar ás $30 \mu g/m^3$. O intervalo con maior número de datos é común ás dúas cidades, o de 5 a $10 \mu g/m^3$, cun máximo en Vigo de 91 días e de 76 en Valladolid.

5.2.3 TERCEIRA VAGA

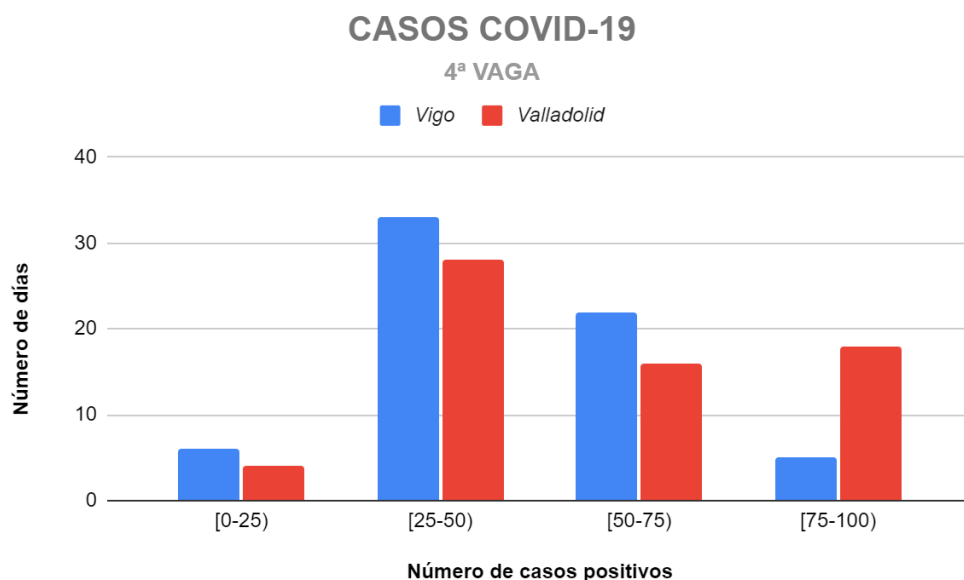


Nesta quinta gráfica representáanse os casos de COVID-19 na terceira vaga. Os datos están separados por intervalos cunha amplitude de 50 casos, sendo o primeiro ([0-50)) o que máis días nos que hai ese número de casos ten acumulados, co máximo de Valladolid con 31 días. No seguinte intervalo ([50-100)) atópase o máximo de Vigo con 31 días.

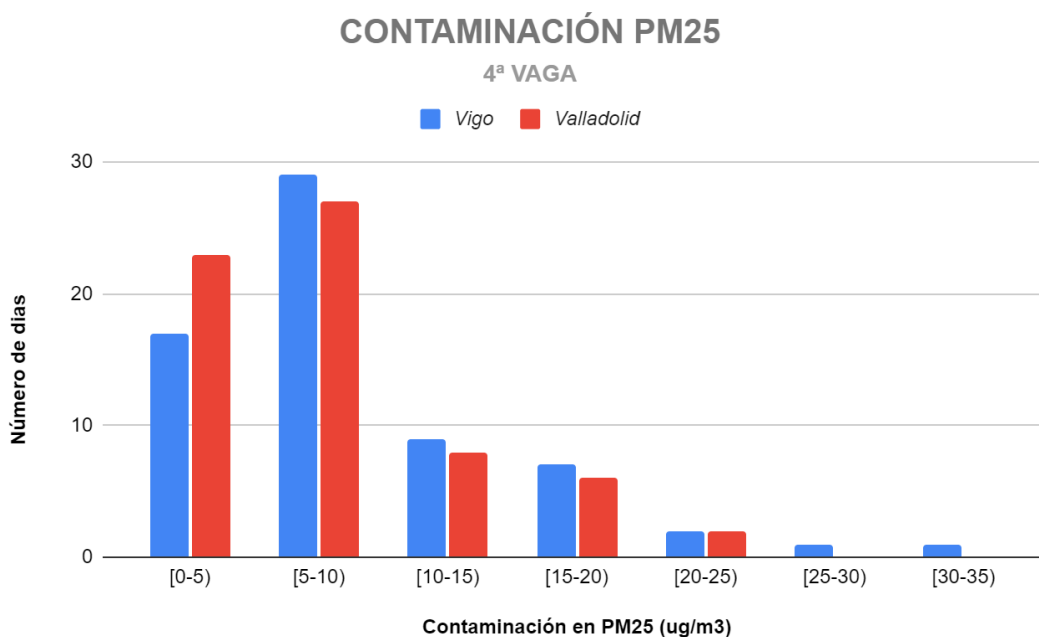


Nesta sexta gráfica representábase a contaminación en partículas PM25 na terceira vaga, os datos están en intervalos de amplitude 5, sendo o primeiro destes ([0-5)) no que se atopa o máximo de Valladolid con 31 días, pola contra, no seguinte intervalo ([5-10)) atópase o máximo de Vigo que é de 40 días.

5.2.4 CUARTA VAGA

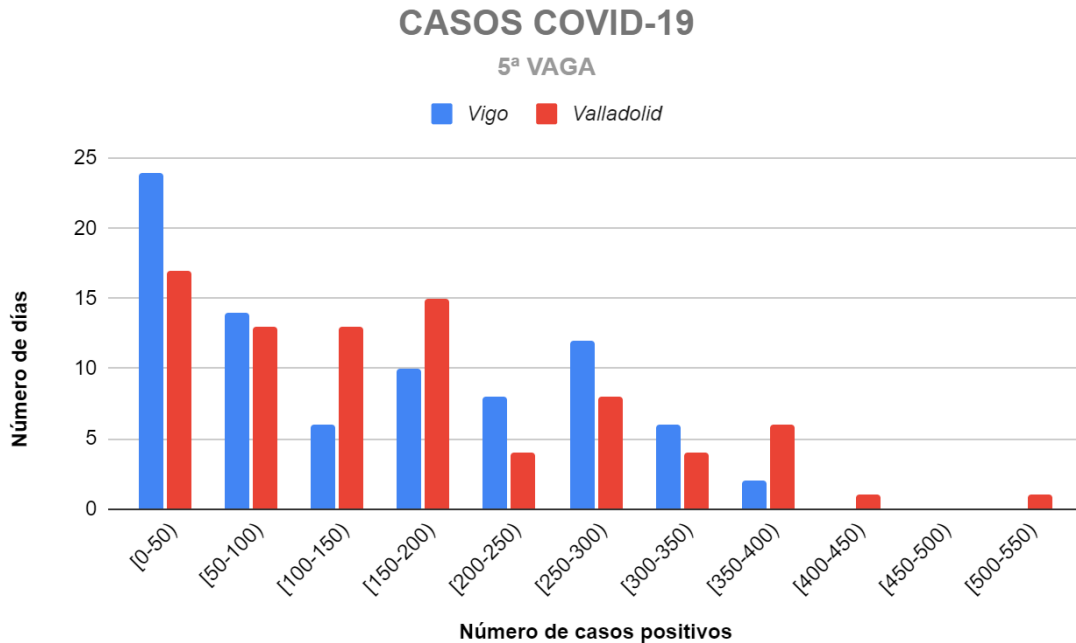


Nesta sétima gráfica representáanse os casos de COVID-19 no transcurso da cuarta vaga, os datos están divididos en 4 intervalos de amplitude 25, no segundo intervalo ([25-50)) atópase o máximo de Vigo con 28 días e o de Valladolid con 33 días.

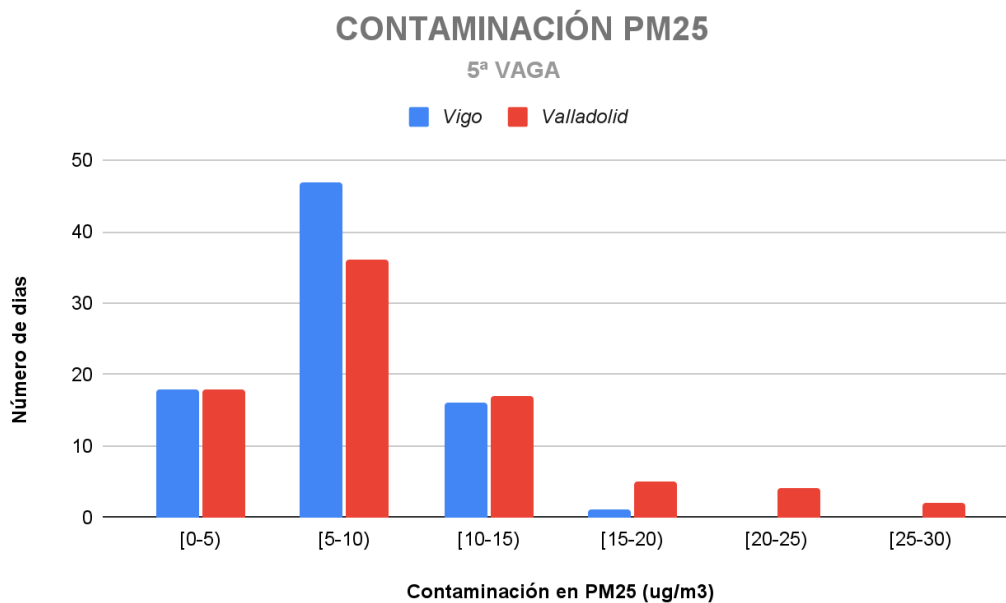


Nesta oitava gráfica representábase a contaminación en PM25 na cuarta vaga, os datos están divididos en sete intervalos de amplitude 5, o máximo de Vigo con 38 días e o de Valladolid con 36 días atópanse no mesmo intervalo ([5-10)).

5.2.5 QUINTA VAGA

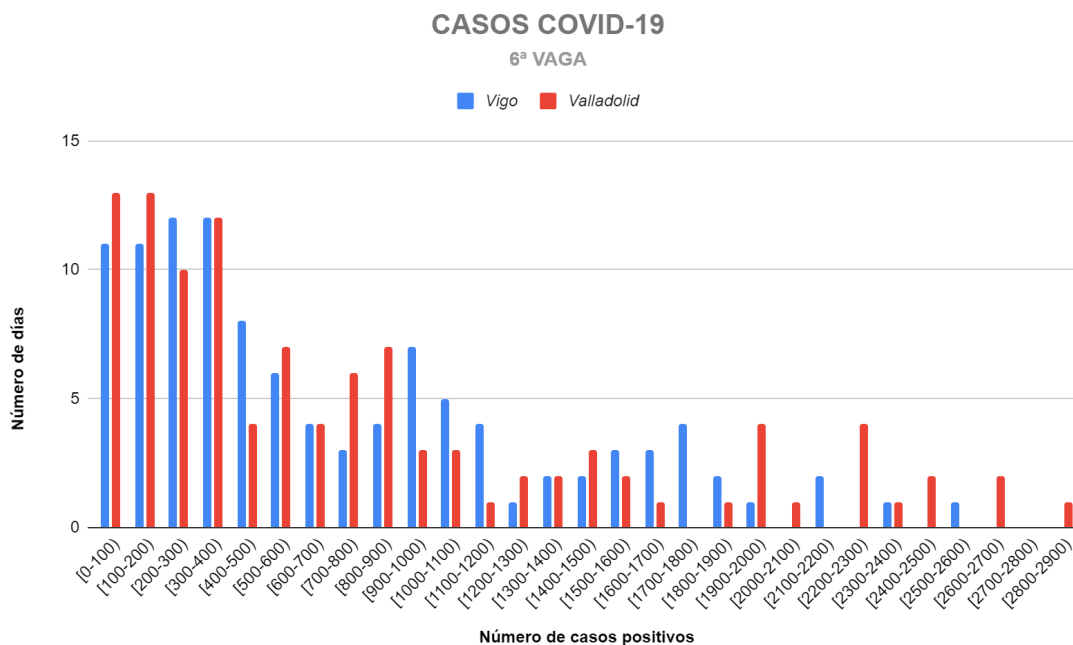


Na novena gráfica representáanse os casos de COVID-19 durante a quinta vaga. Os datos están amosados en intervalos de amplitude 50, o primeiro ([0-50)) será o intervalo no que se atope o máximo de Vigo (24) e Valladolid (16).

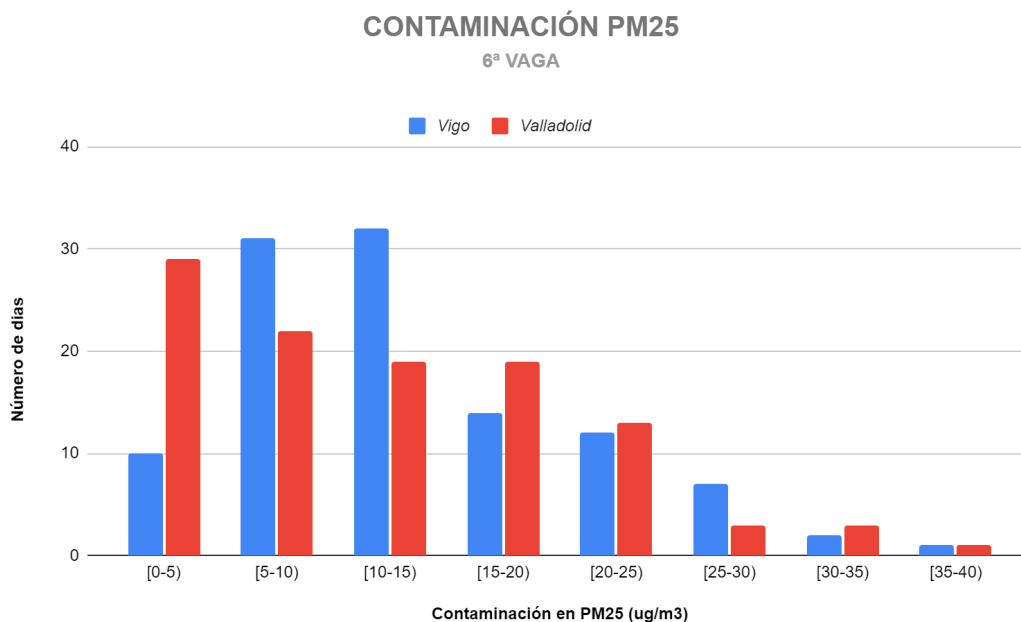


Nesta décima gráfica representáanse a contaminación en PM25 durante a quinta vaga. Os datos están divididos por intervalos de amplitude 5, sendo o segundo o que máis datos de contaminación ten acumulados, co máximo de Vigo (47) e o máximo de Valladolid (36).

5.2.6 SEXTA VAGA



Nesta undécima gráfica represéntase os casos de COVID-19 durante a sexta ola. Os intervalos teñen amplitude de 100, no terceiro e no cuarto intervalo ([200-300) e [300-400)) atopamos o máximo de datos acumulados de Vigo (12) mentres que o de Valladolid (13) atópase nos dous primeiros intervalos ([0-100) e [100-200)).



Nesta última gráfica represéntase a contaminación en PM25 durante a sexta vaga. Neste caso os intervalos teñen amplitude 5. O máximo de datos de Vigo (32) atópase no terceiro intervalo ([10-15)) mentres que o de Valladolid (29) atópase no primeiro intervalo ([0-5)).

5.3 TÁBOAS DE ANÁLISE DAS MEDIDAS DE CENTRALIZACIÓN E DISPERSIÓN

Neste apartado do informe recolleremos os valores das medidas de centralización e de dispersión calculadas en dúas táboas diferentes (unha para cada variable) onde incluímos as dúas cidades obxecto de estudo e as respectivas vagas pandémicas.

5.3.1 CASOS POSITIVOS DE COVID-19

	1ª VAGA		2ª VAGA		3ª VAGA		4ª VAGA		5ª VAGA		6ª VAGA			
	Vigo	Valladolid	Vigo	Valladolid	Vigo	Valladolid	Vigo	Valladolid	Vigo	Valladolid	Vigo		Valladolid	
MEDIA	27,19	57,81	66,02	198,44	114,92	187,90	47,35	55,68	146,34	161,59	727,98		784,86	
MEDIANA	12,5	37,5	37,5	125	87,5	75	37,5	62,5	112,5	125	550		550	
MODA	12,5	12,5	12,5	75	37,5	25	37,5	37,5	37,5	25	250	350	50	150
DT	28,24	49,79	61,46	175,67	86,91	211,88	17,40	18,27	110,65	119,28	608,45		736,95	
CO.VA	1,04	0,86	0,93	0,89	0,76	1,13	0,37	0,33	0,76	0,74	0,84		0,94	
CO. CORREL	0,34		0,63		0,91		0,46		0,70		0,74			

Nesta táboa represéntanse as medidas de centralización e dispersión dos casos de COVID-19 durante a pandemia nas dúas cidades obxecto de estudo. Na primeira vaga, a media de casos diarios en Vigo foi de 27,19 mentres, que en Valladolid foi de 57,81. Se miramos a mediana, vemos que a de Valladolid (37,5) é superior á de Vigo (12,5), isto implica que en Valladolid o 50% dos días da primeira vaga houbo menos de 37,5 casos e que no outro 50% houbo máis de 37,5, mentres que en Vigo houbo o 50% dos días menos de 12,5 casos e no outro 50% máis de 12,5. Por outra banda, as dúas cidades comparten a moda de 12,5, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 49,79, tendo Vigo 28,24, despois, o coeficiente de variación, no que, neste caso, é maior en Vigo cun 1,04, mentres que en Valladolid é de 0,86, por último temos o coeficiente de correlación, que é de 0,34 .

Na segunda vaga, a media de casos diarios en Vigo foi de 66,02 mentres que en Valladolid foi de 198,44. Se miramos a mediana, vemos que a de Vigo (37,5) é inferior á de Valladolid (125), isto implica que en Valladolid o 50% dos días da segunda vaga houbo menos de 125 casos e que no outro 50% houbo máis de 125, mentres que en Vigo houbo o 50% dos días menos de 37,5 casos e no outro 50% máis de 37,5. Por outra

banda, na moda, é inferior Vigo con 12,5, mentres que Valladolid ten un 75, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 175,67, tendo Vigo 61,46, despois, no cociente de variación é maior en Vigo cun 0,93, mentres que en Valladolid é de 0,89, por último temos o coeficiente de correlación, que é de 0,63.

Na terceira vaga, a media de casos diarios en Vigo foi de 114,92 mentres, que en Valladolid foi de 187,90. Se miramos a mediana, vemos que a de Vigo (87,5) é superior á de Valladolid (25), isto implica que en Vigo o 50% dos días da terceira vaga houbo menos de 87,5 casos e que no outro 50% houbo máis de 87,5, mentres que en Valladolid houbo o 50% dos días menos de 25 casos e no outro 50% máis de 25. Por outra banda, na moda, é superior Vigo cun 37,5, mentres que Valladolid ten un 25, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 211,88, tendo Vigo 86,91, despois, vén o cociente de variación, no que, neste caso, é inferior en Vigo cun 0,76, mentres que en Valladolid é de 1,13, por último temos o coeficiente de correlación, que é de 0,91.

Na cuarta vaga, a media de casos diarios en Vigo foi de 47,35 mentres, que en Valladolid foi de 55,68. Se miramos a mediana, vemos que a de Vigo (37,5) é inferior á de Valladolid (62,5), isto implica que en Valladolid o 50% dos días da cuarta vaga houbo menos de 62,5 casos e que no outro 50% houbo máis de 62,5, mentres que en Vigo houbo o 50% dos días menos de 37,5 casos e no outro 50% máis de 37,5. Por outra banda, na moda, Vigo é igual Valladolid cun 37,5, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 18,27, tendo Vigo 17,40, despois, vén o cociente de variación, no que, neste caso, é maior en Vigo con 0,37, mentres que en Valladolid é de 0,33, por último temos o coeficiente de correlación, que é de 0,46.

Na quinta vaga, a media de casos diarios en Vigo foi de 146,34 mentres que en Valladolid foi de 16,59. Se miramos a mediana, vemos que a de Vigo (112,5) é inferior á de Valladolid (125), isto implica que en Valladolid o 50% dos días da quinta vaga houbo menos de 125 casos e que no outro 50% houbo máis de 125, mentres que en Vigo houbo o 50% dos días menos de 112,5 casos e no outro 50% máis de 112,5. Por outra banda, na moda, é superior Vigo cun 37,5, mentres que Valladolid ten un 25, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 119,28, tendo Vigo 110,65, despois, vén o cociente de variación, no que, neste caso, é maior en Vigo cun 0,76, mentres que en Valladolid é de 0,74, por último temos o coeficiente de correlación, que é de 0,70.

Na sexta vaga, a media de casos diarios en Vigo foi de 727,98 mentres que en Valladolid foi de 784,86. Se miramos a mediana, vemos que a de Vigo (550) é igual á de Valladolid (550), isto implica que nas dúas cidades durante a sexta vaga houbo o 50% dos días menos de 550 casos e no outro 50% máis de 550. Por outra banda, na moda, é superior Vigo con dúas modas 250 e 350, mentres que Valladolid son de 50 e 150, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 736,5, tendo Vigo 608,45, despois, vén o cociente de variación, no que, neste caso, é inferior en Vigo cun 0,70, mentres que en Valladolid é de 0,74, por último temos o coeficiente de correlación, que é de 0,74.

5.3.2 CONTAMINACIÓN DO AIRE EN PM25

	1ª VAGA		2ª VAGA		3ª VAGA		4ª VAGA		5ª VAGA		6ª VAGA	
	Vigo	Valladolid	Vigo	Valladolid	Vigo	Valladolid	Vigo	Valladolid	Vigo	Valladolid	Vigo	Valladolid
MEDIA	9,06	8,56	8,40	10,27	11,69	9,54	9,09	7,73	7,5	9,27	13,46	12,09
MEDIANA	7,5	7,5	7,5	7,5	7,5	7,5	7,5	7,5	7,5	7,5	12,5	12,5
MODA	7,5	7,5	7,5	7,5	7,5	2,5	7,5	7,5	7,5	7,5	12,5	7,5
DT	3,95	4,94	5,06	6,37	7,87	7,96	6,23	4,98	3,13	5,97	7,29	8,51
CO.VA	0,44	0,57	0,60	0,62	0,67	0,83	0,69	0,64	0,42	0,64	0,54	0,70
CO. CORREL	0,45		0,39		0,52		0,49		0,30		0,49	

Nesta táboa representáanse as medidas de centralización e dispersión da contaminación en PM25 durante a pandemia nas dúas cidades obxecto de estudo. Na primeira vaga, a media de contaminación diaria en Vigo foi de 9,06 mentres que en Valladolid, foi de 8,56. Se miramos a mediana, vemos que a de Valladolid (7,5) é igual cá a de Vigo, isto implica que nas dúas cidades durante a primeira vaga houbo o 50% dos días menos de $7,5 \mu\text{g}/\text{m}^3$ de contaminación en PM25 e no outro 50% máis de $7,5 \mu\text{g}/\text{m}^3$. Por outra banda, as dúas cidades comparten a moda de 7,5, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 4,94, tendo Vigo 3,95, despois, o coeficiente de variación, no que, neste caso, é maior en Valladolid cun 0,57, mentres que en Vigo é de 0,44, por último temos o coeficiente de correlación, que é de 0,45.

Na segunda vaga, a media de contaminación diaria en Vigo foi de 8,40 mentres que en Valladolid, foi de 10,27. Se miramos a mediana, vemos que a de Valladolid (7,5) é igual cá a de Vigo, isto implica que nas dúas cidades durante a segunda vaga houbo o 50% dos días menos de $7,5 \mu\text{g}/\text{m}^3$ de contaminación en PM25 e no outro 50% máis de $7,5 \mu\text{g}/\text{m}^3$. Por outra banda, as dúas cidades comparten a moda de 7,5, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 6,37, tendo Vigo 5,06, despois, o coeficiente de variación, no que, neste caso, é maior en Valladolid cun 0,62, mentres que en Vigo é de 0,60, por último temos o coeficiente de correlación, que é de 0,39.

Na terceira vaga, a media de contaminación diaria en Vigo foi de 11,69 mentres que en Valladolid, foi de 9,54. Se miramos a mediana, vemos que a de Valladolid (7,5) é igual cá a de Vigo, isto implica que nas dúas cidades durante a terceira vaga houbo o 50% dos días menos de 7,5

$\mu\text{g}/\text{m}^3$ de contaminación en PM25 e no outro 50% máis de $7,5 \mu\text{g}/\text{m}^3$. Por outra banda, as dúas cidades comparte a moda de 7,5, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 7,96, tendo Vigo 7,87, despois, o coeficiente de variación, no que, neste caso, é maior en Valladolid cun 0,83, mentres que en Vigo é de 0,67, por último temos o coeficiente de correlación, que é de 0,52.

Na cuarta vaga, a media de contaminación diaria en Vigo foi de 7,5 mentres que en Valladolid, foi de 9,27. Se miramos a mediana, vemos que a de Valladolid (7,5) é igual cá a de Vigo, isto implica que nas dúas cidades durante a cuarta vaga houbo o 50% dos días menos de $7,5 \mu\text{g}/\text{m}^3$ de contaminación en PM25 e no outro 50% máis de $7,5 \mu\text{g}/\text{m}^3$. Por outra banda, as dúas cidades comparten a moda de 7,5, na seguinte fila, na desviación típica, Vigo supera a Valladolid cun 6,23, tendo Valladolid 4,98, despois, o coeficiente de variación, no que, neste caso, é menor en Valladolid cun 0,64, mentres que en Vigo é de 0,69, por último temos o coeficiente de correlación, que é de 0,49.

Na quinta vaga, a media de contaminación diaria en Vigo foi de 9,09 mentres que en Valladolid, foi de 7,73. Se miramos a mediana, vemos que a de Valladolid (7,5) é igual cá a de Vigo, isto implica que nas dúas cidades durante a quinta vaga houbo o 50% dos días menos de $7,5 \mu\text{g}/\text{m}^3$ de contaminación en PM25 e no outro 50% máis de $7,5 \mu\text{g}/\text{m}^3$. Por outra banda, as dúas cidades comparten a moda de 7,5, na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 5,97, tendo Vigo 3,13 despois, o coeficiente de variación, no que, neste caso, é maior en Valladolid cun 0,64, mentres que en Vigo é de 0,42, por último temos o coeficiente de correlación, que é de 0,30 .

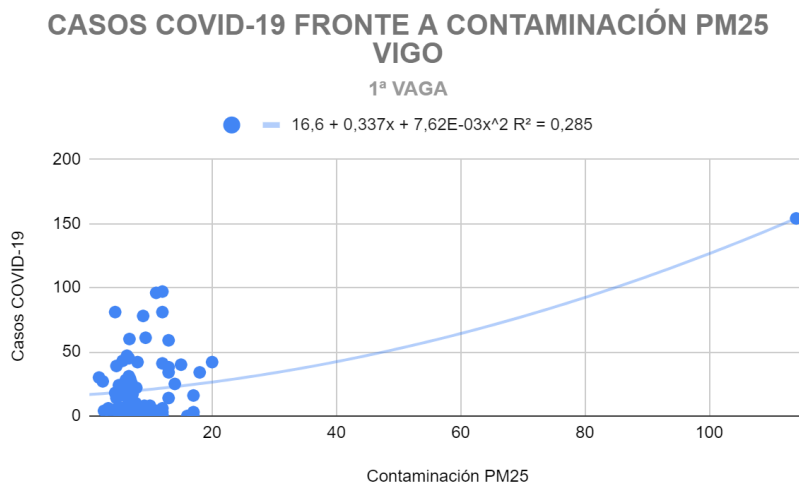
Na sexta vaga, a media de contaminación diaria en Vigo foi de 13,46 mentres, que en Valladolid, foi de 12,09 Se miramos a mediana, vemos que a de Valladolid (12,5) é igual cá a de Vigo, isto implica que nas dúas cidades durante a sexta vaga houbo o 50% dos días menos de $12,5 \mu\text{g}/\text{m}^3$ de contaminación en PM25 e no outro 50% máis de $12,5 \mu\text{g}/\text{m}^3$. Por outra banda, a moda de Vigo (12,5) é maior á de Valladolid (7,5), na seguinte fila, na desviación típica, Valladolid supera a Vigo cun 8,51 tendo Vigo 7,29 despois, o coeficiente de variación, no que, neste caso, é maior en Valladolid cun 0,70, mentres que en Vigo é de 0,54, por último temos o coeficiente de correlación, que é de 0,49 .

5.4 ESTUDO DE CORRELACIÓN

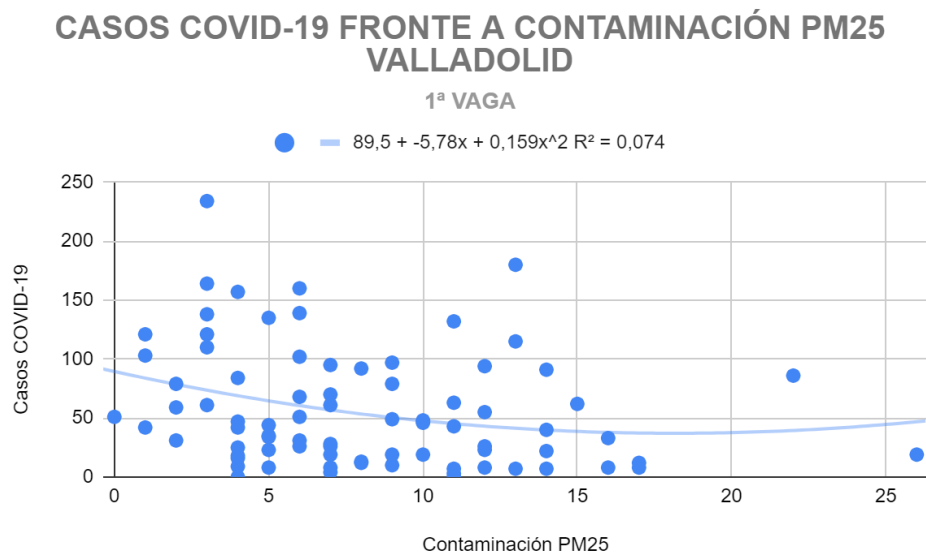
Para realizar o estudo de correlación realizamos gráficas de dispersión. Nestas gráficas incluímos no eixo X os datos de contaminación e no eixo Y os datos de COVID, para que o estudo se realizase de xeito que os casos de COVID-19 dependesen da contaminación.

Nestas gráficas incluímos liñas de tendencia máis axeitadas á dispersión que nos permiten coñecer a correlación entre as variables extraendo o R^2 que indica o preto que están os datos da regresión lineal e que nos permite axustar as relacións e coñecer se as hay.

5.4.1 PRIMEIRA VAGA



Nesta gráfica incluímos a nube de datos da primeira vaga en Vigo onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,285.

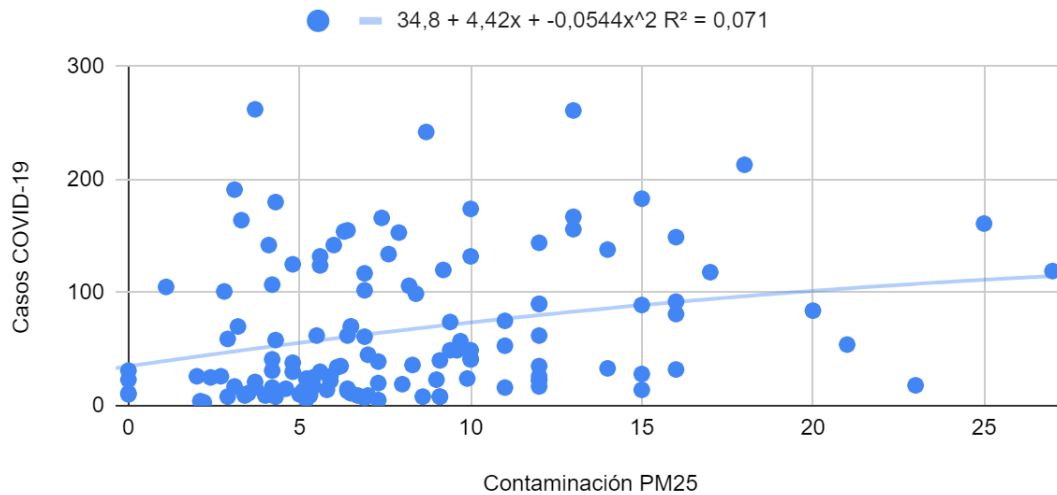


Nesta gráfica incluímos a nube de datos da primeira vaga en Valladolid onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,074.

5.4.2 SEGUNDA VAGA

CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VIGO

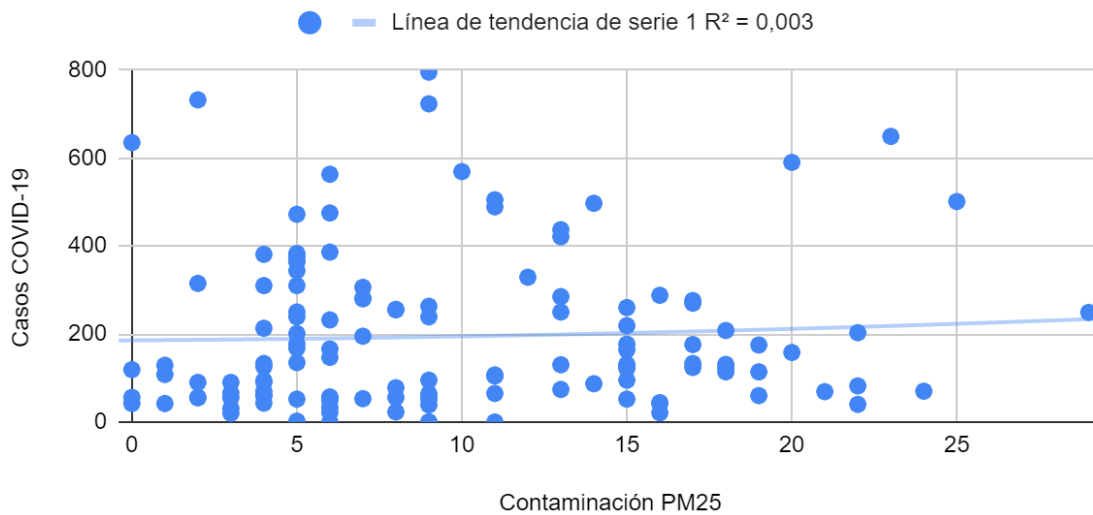
2ª VAGA



Nesta gráfica incluímos a nube de datos da segunda vaga en Vigo onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,071.

CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VALLADOLID

2ª VAGA

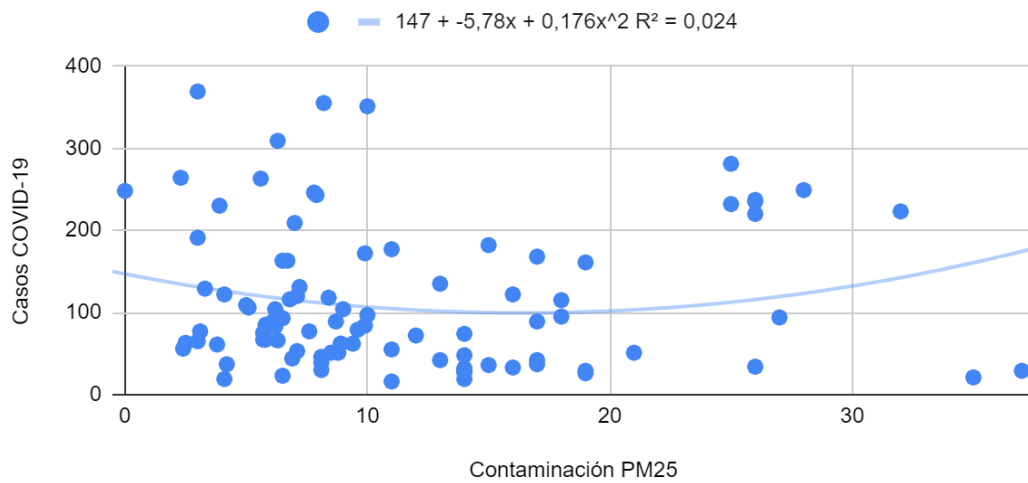


Nesta gráfica incluímos a nube de datos da segunda vaga en Valladolid onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,003.

5.4.3 TERCEIRA VAGA

CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VIGO

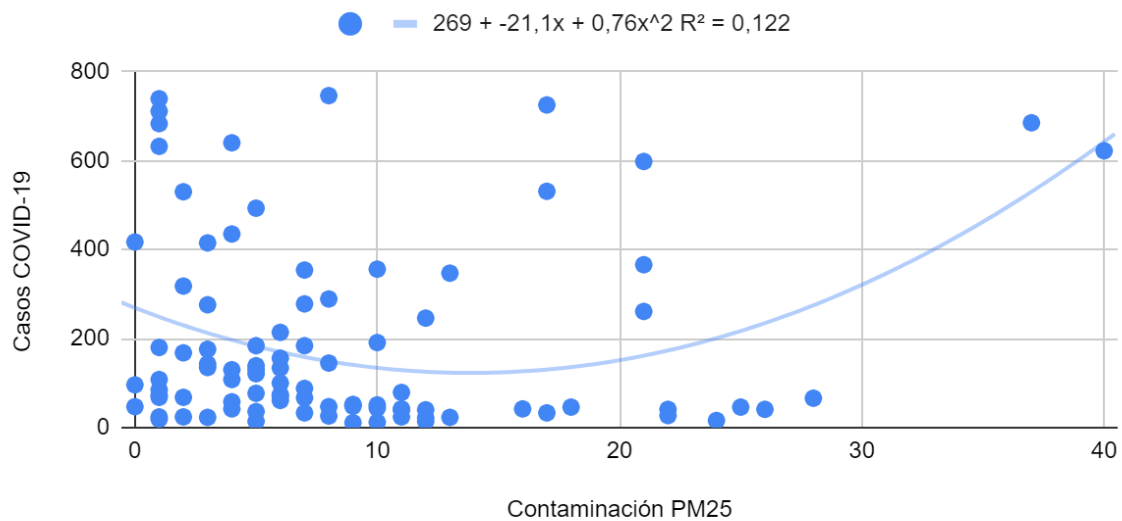
3ª VAGA



Nesta gráfica incluímos a nube de datos da terceira vaga en Vigo onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,024.

CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VALLADOLID

3ª VAGA

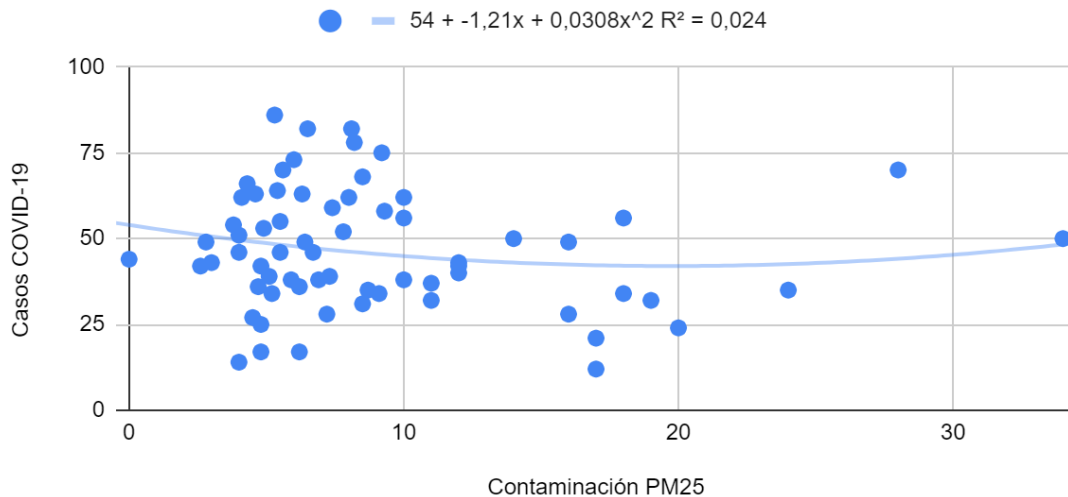


Nesta gráfica incluímos a nube de datos da terceira vaga en Valladolid onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,122.

5.4.4 CUARTA VAGA

CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VIGO

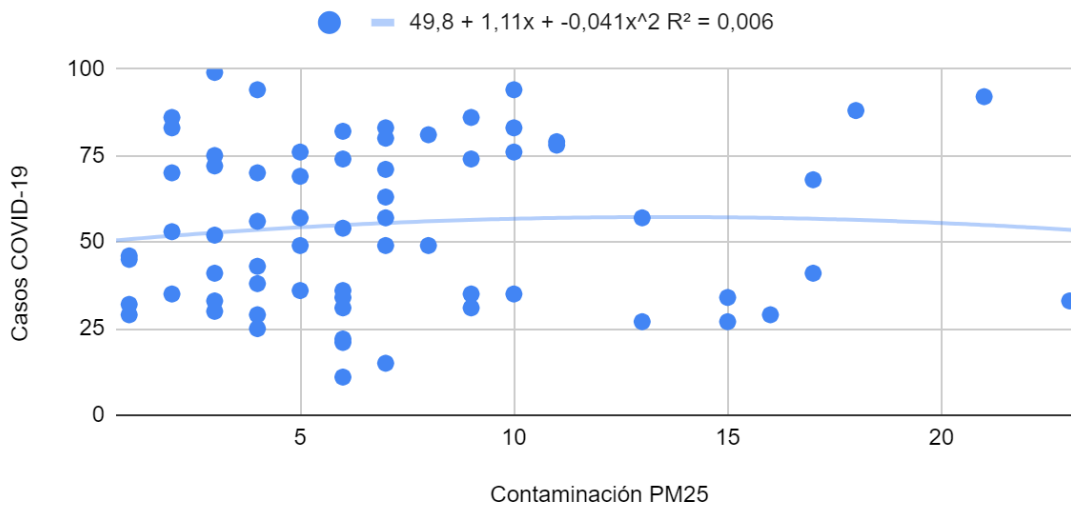
4ª VAGA



Nesta gráfica incluímos a nube de datos da cuarta vaga en Vigo onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,024.

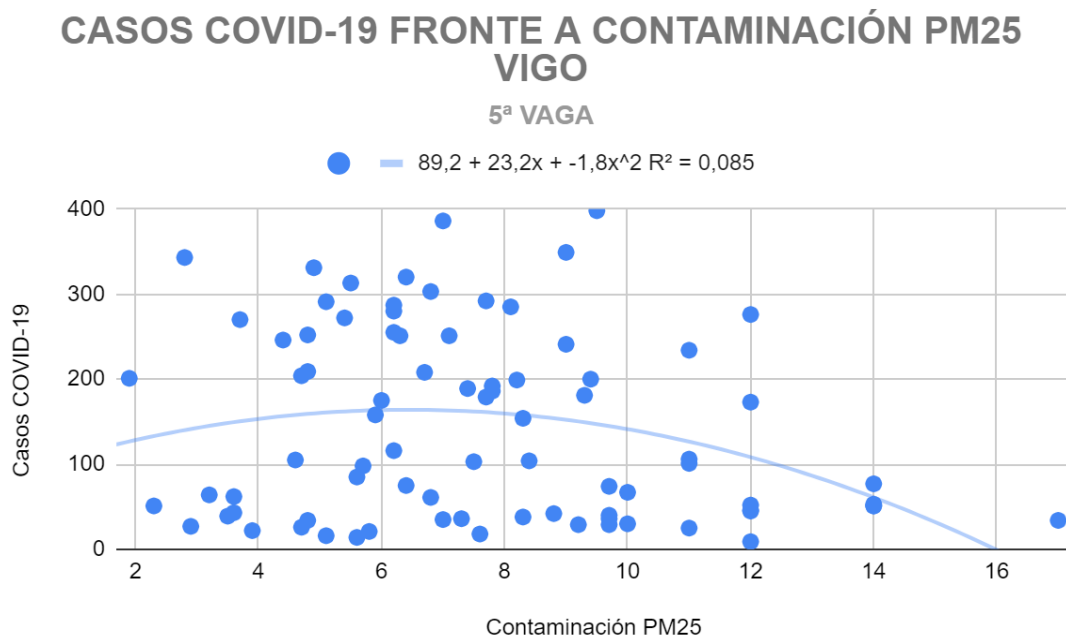
CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VALLADOLID

4ª VAGA

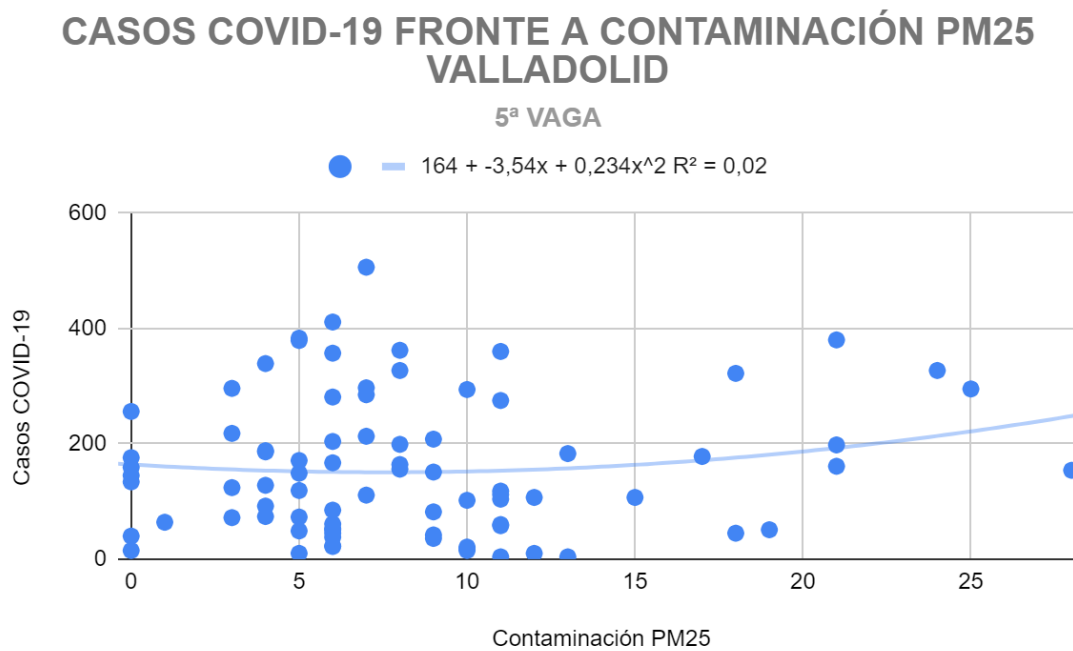


Nesta gráfica incluímos a nube de datos da cuarta vaga en Valladolid onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,006.

5.4.5 QUINTA VAGA



Nesta gráfica incluímos a nube de datos da quinta vaga en Vigo onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,085.

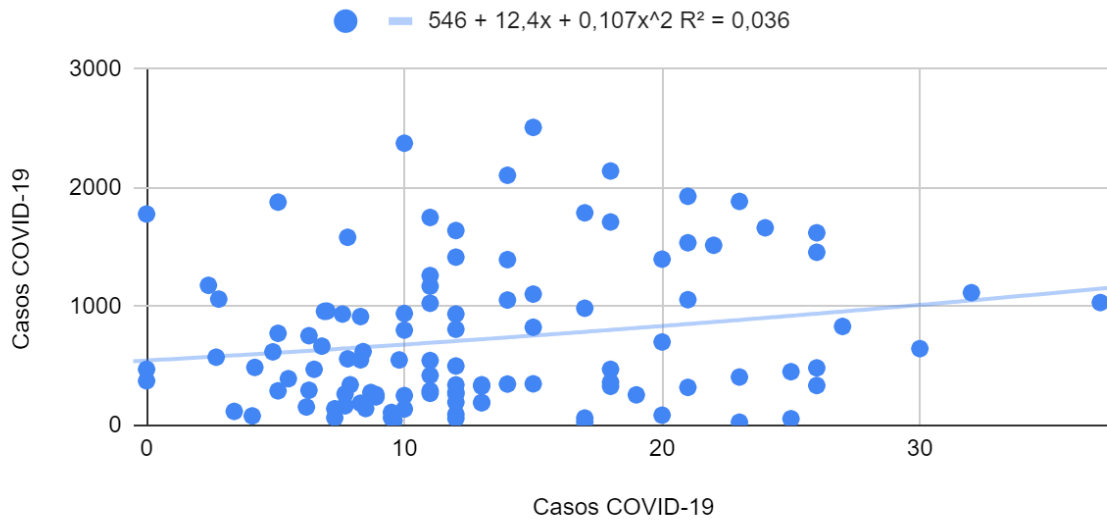


Nesta gráfica incluímos a nube de datos da quinta vaga en Valladolid onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,02.

5.4.6 SEXTA VAGA

CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VIGO

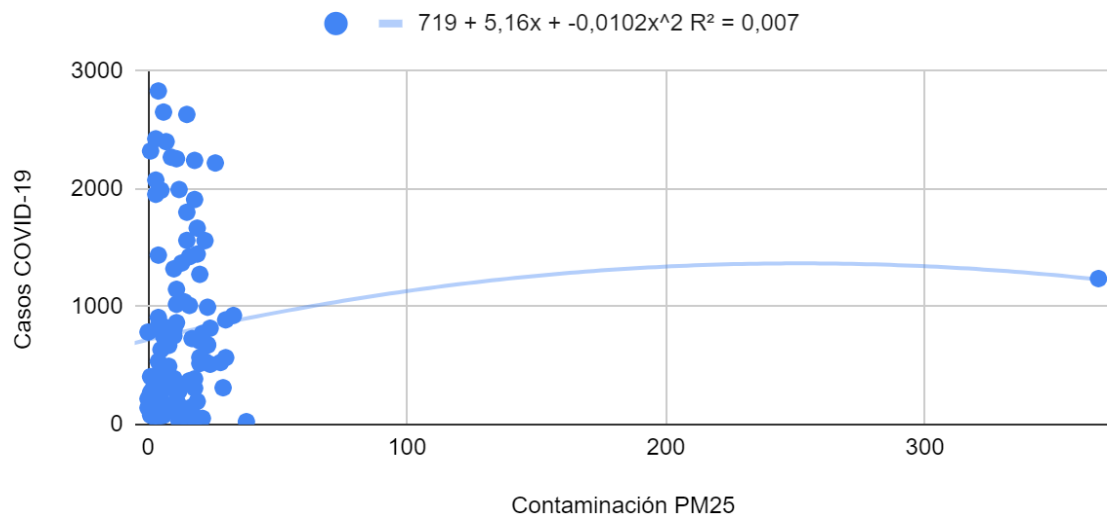
6ª VAGA



Nesta gráfica incluímos a nube de datos da sexta vaga en Vigo onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,036.

CASOS COVID-19 FRONTE A CONTAMINACIÓN PM25 VALLADOLID

6ª VAGA



Nesta gráfica incluímos a nube de datos da sexta vaga en Valladolid onde a liña de tendencia máis axeitada trátase dunha función polinómica de segundo grao. Neste caso o R^2 ten como valor 0,007.

6. CONCLUSIONES

Entrámbalas dúas cidades observamos que Valladolid ten un maior número de casos de COVID-19 e de contaminación que en Vigo polo que podíamos afirmar que a contaminación do aire influía no número de casos de coronavirus, pero esta relación tamén podería vir dada polo maior número de habitantes que ten Valladolid.

Outras conclusións que podíamos extraer teñen que ver coa sexta vaga. Sabemos que a sexta vaga foi a onda pandémica máis explosiva de todas onde se superaron todos os récords de contaxios. Pero esta explosión de casos, que estivo influída pola aparición da variante Ómicron, tamén viu acompañada dun aumento da contaminación do aire, de feito esta vaga foi na que se rexistraron maiores niveis de contaminación.

Pero unha vez realizado o estudo de correlación concluímos que o noso estudo non é concluínte, xa que non podemos afirmar nin desmentir que se poidan establecer relacións entre a contaminación do aire e os casos de COVID-19, polo que a nosa hipótese inicial aínda non pode ser confirmada ou rexeitada.

Esta conclusión dáse xa que os nosos R^2 están moi preto do valor de 0 polo que non podemos establecer a relación xa que cando o R^2 tende a 0 as variables están menos relacionadas, mentres ue cando tende a 1 están máis relacionadas e se poden establecer relacións funcionais.

7. PROPOSTAS DE MELLORA

Para a mellora o noso estudo a principal proposta que propoñemos de cara a próximas investigación cuxa base sexa este traballo de investigación sería a inclusión doutras enfermidades respiratorias, como por exemplo a gripe ou as pneumonías, a parte a COVID; para estudar se se poden establecer relacións entre elas e a contaminación. Resumidamente a nosa proposta sería a de aplicar esta mesma investigación a outras enfermidades de carácter respiratorio.

Outra proposta de mellora que nos gustaría que se aplicase de cara a investigacións futuras sería a de centralizar os nosos datos para realizar unha investigación moito máis centralizada para concluír definitivamente se se poden establecer relacións entre estas dúas variables do estudo.

8. REFERENCIAS

¹ Elena Salamanca-Fernández, Miguel Rodríguez Barranco y María José Sánchez, 2021. Influencia de la temperatura ambiental y la contaminación en la transmisión del SARS-CoV-2. *Revista Española de Salud Pública*. Madrid: 95: 20 de enero e202101003