

VISUALIZACIÓN DE DATOS

PRÁCTICA 2

Carlos Mas Estévez

1.- Objetivos

El conjunto de datos utilizado ha sido obtenido del estudio n.º 3325 del CIS, titulado *Encuesta sobre relaciones sociales y afectivas en tiempos de pandemia de la Covid-19* y realizada del 27 de mayo al 17 de junio de 2021 a toda la población española mayor de 18 años de manera proporcional según la región y localidad. Se puede consultar desde el siguiente enlace:

https://www.cis.es/cis/opencm/ES/1_encuestas/estudios/ver.jsp?estudio=14571

En él, se recogen 2972 registros de respuestas de la gente sobre sus situaciones sociales y afectivas. Con esta información, podremos estudiar multitud de situaciones y reponernos las siguientes preguntas:

[1] - ¿Tener una relación puede afectar a la percepción que se tiene del uso de internet?

[2] - ¿La satisfacción en la vida familiar afecta a la satisfacción en distintos campos relacionados a la vida social o afectiva?

[3] - ¿Cómo son de satisfactorias las relaciones afectivas según el nivel de estudios de cada componente de la pareja? ¿y según a su clase social?

Así pues, centraremos nuestro proyecto de visualización a la resolución de las tres preguntas efectuadas anteriormente.

Los archivos necesarios para la ejecución de las visualizaciones y los archivos .csv necesarios se encuentran en el siguiente repositorio GitHub:

https://github.com/masescar/PRA2_visualizacion_datos

Todas las visualizaciones que son explicadas a continuación han sido recopiladas en formato presentación en Flourish, accesible desde el siguiente link:

<https://public.flourish.studio/story/1808448/>

2.- Resolución de las preguntas

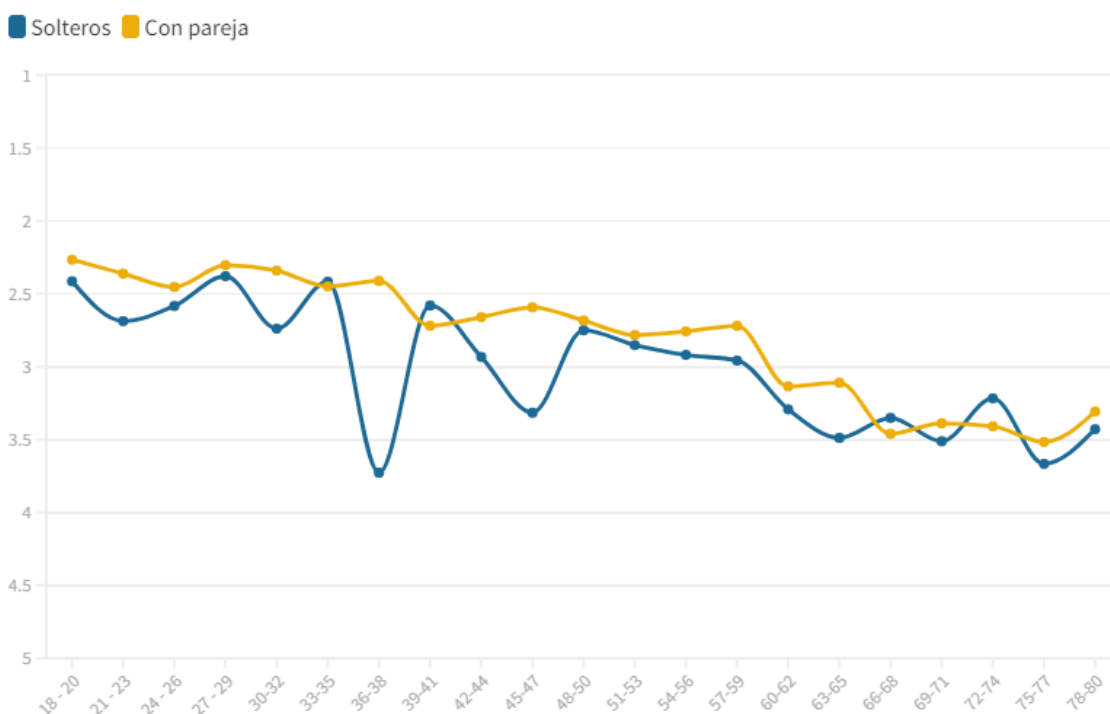
2.1.- Percepción del uso de internet

Para responder a esta pregunta y visualizarla de manera correcta, hemos decidido agrupar la media de afinidad sobre el uso de internet para conocer gente (puntuada del 1 al 5, de mayor a menor afinidad) según la tenencia de pareja y agrupados según la edad (agrupados de 3 en 3 años, empezando desde los 18-20 y acabando en los 90-92).

Hemos realizado una visualización de tipo gráfico de puntos unidos mediante líneas suavizadas con el fin de comprobar diferencias en la tendencia sobre la afinidad del uso de internet según la edad y la tenencia de pareja.

Opinión sobre uso de internet para conocer gente

1 muy bien / 2 bien / 3 neutral / 4 mal / 5 muy mal



Se puede consultar la visualización en el siguiente link:

<https://public.flourish.studio/visualisation/12515107/>

Se observa que la gente que tiene pareja tiene una opinión, por lo general, ligeramente mejor sobre el uso de internet respecto a la gente soltera, teniendo ambas líneas una tendencia bajista a medida que se aumenta la edad. Las dos líneas tienen colores complementarios para destacar más entre ellas.

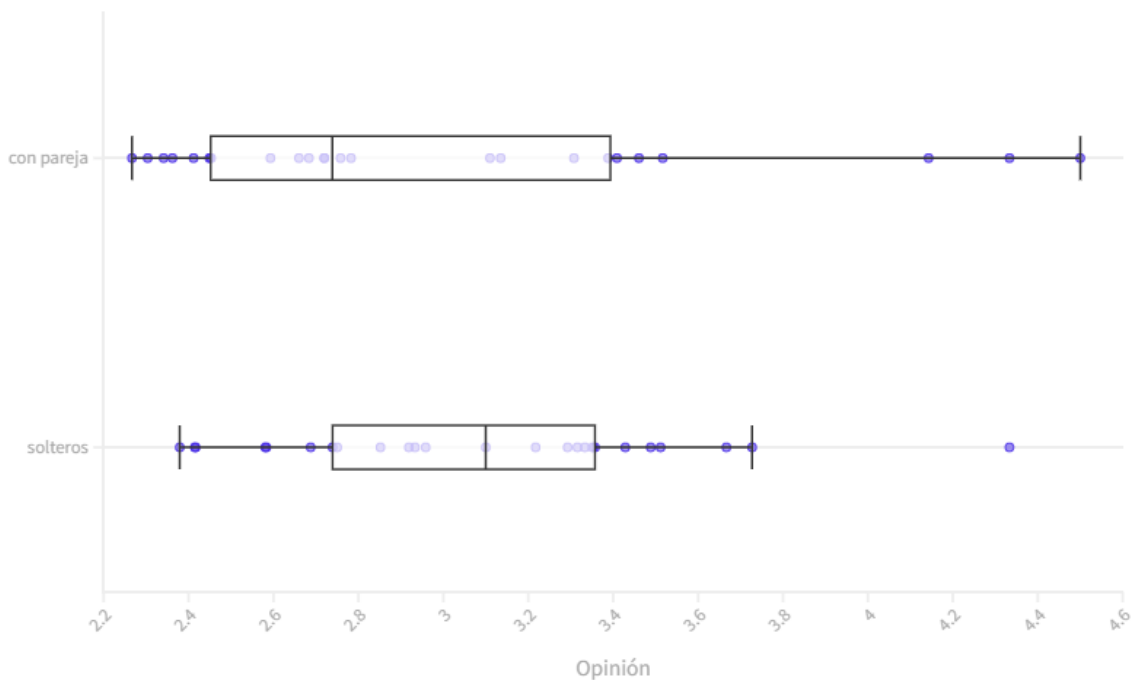
Para reforzar nuestras conclusiones, además hemos realizado otra visualización al respecto que consiste en un gráfico boxplot para determinar la distribución de las opiniones de las personas sobre el uso de internet según la tenencia de pareja (medias agrupadas por franjas de edad).

Se puede consultar la visualización Boxplot en el siguiente link:

<https://public.flourish.studio/visualisation/12515346/>

Distribución de la opinión sobre uso de internet

1 - Buena opinión / 5 - Mala opinión



Observamos que la mediana de la opinión sobre el uso de internet por parte de los solteros tiene un valor mayor respecto a la de las personas con pareja, lo cual es coherente con el gráfico anterior y nos permite sentenciar que las personas solteras tienen una peor opinión sobre el uso de internet para conocer gente respecto a las personas que tienen una relación.

2.2.- Satisfacción en la vida familiar según la vida social y afectiva

Para contestar a esta pregunta, hemos optado por realizar una visualización scatter plot con el fin de representar el comportamiento de una variable (satisfacción en ámbito familiar, puntuada del 1 al 10 en orden ascendente) en función de otras 2 representadas en los ejes x, e y (satisfacción afectiva y social, respectivamente y puntuada de forma idéntica).

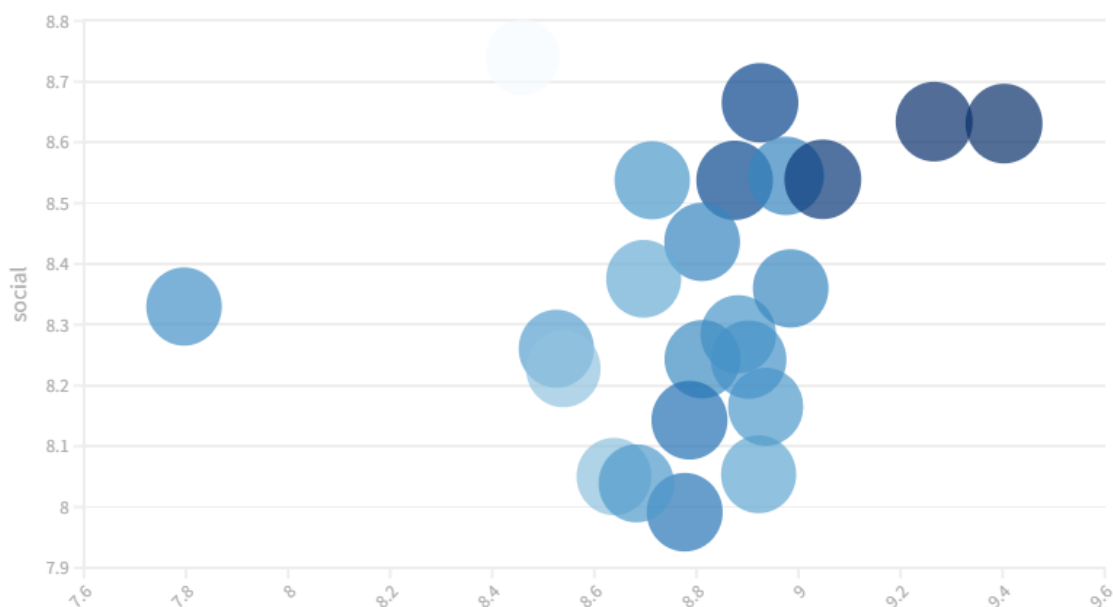
Podemos consultar la visualización en el siguiente link:

<https://public.flourish.studio/visualisation/12515925/>

Felicidad en entorno familiar según la felicidad en entorno social y afectivo

Agrupado por edades de 3 en 3 desde los 18 hasta los 86

familia 8.12  9.02



Cada círculo azul representa la media de satisfacción familiar en un determinado rango de edad (agrupados de 3 en 3 como antes). Según su intensidad en el color, su valor será mayor (más oscuro) o menor (más claro). Observamos que, salvo casos puntuales, hay una tendencia general a adoptar valores más oscuros cuanto mayor valor tengan las variables de satisfacción social y afectiva (ejes y, x respectivamente). Por lo tanto, podríamos inducir que a mayor satisfacción familiar, mayor será la satisfacción del individuo en ámbitos sociales y afectivos.

Consideramos que esta visualización ha sido una buena elección porque nos permite identificar de forma simple el comportamiento de una variable solamente viendo el color que adopta, pues en este caso un color más oscuro implica un valor mayor, y el hecho de que muchos puntos de similar color estén en la misma región del gráfico implica que puede existir una correlación entre la variable de satisfacción familiar (la que le asignamos el color) y las otras dos (situadas en los ejes).

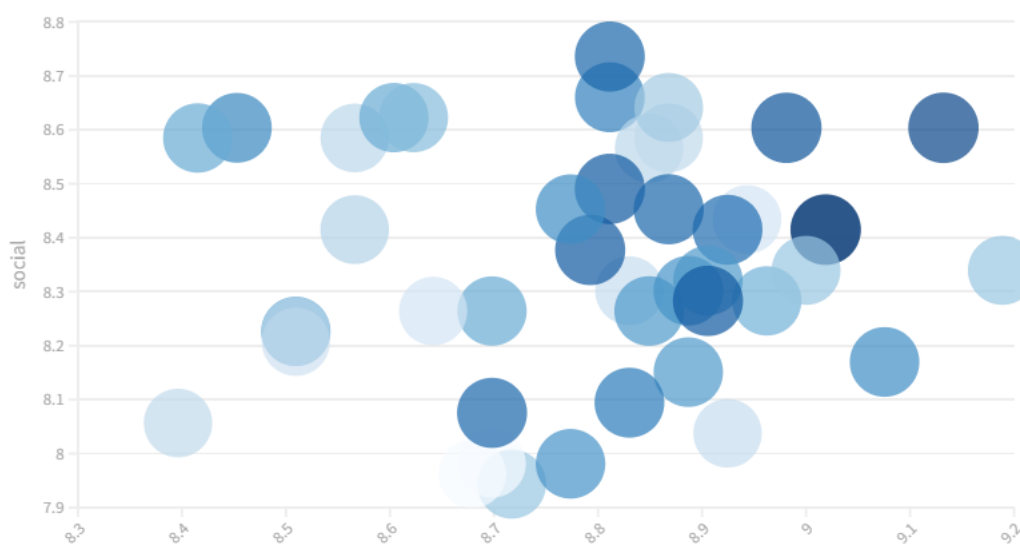
Debido a que podríamos achacar este suceso al comportamiento y situación de los individuos en los distintos rangos de edad, también se ha optado por realizar la misma visualización agrupando al azar 44 grupos de 53 miembros cada uno (para redondear en el número total de registros) a fin de comprobar que la tendencia observada anteriormente es independiente de la edad. Veamos en la siguiente visualización qué obtenemos:

<https://public.flourish.studio/visualisation/12517146/>

Felicidad en entorno familiar según la felicidad en entorno social y afectivo

Agrupado al azar en 44 grupos con 53 muestras cada uno

familia 8.42  9.15



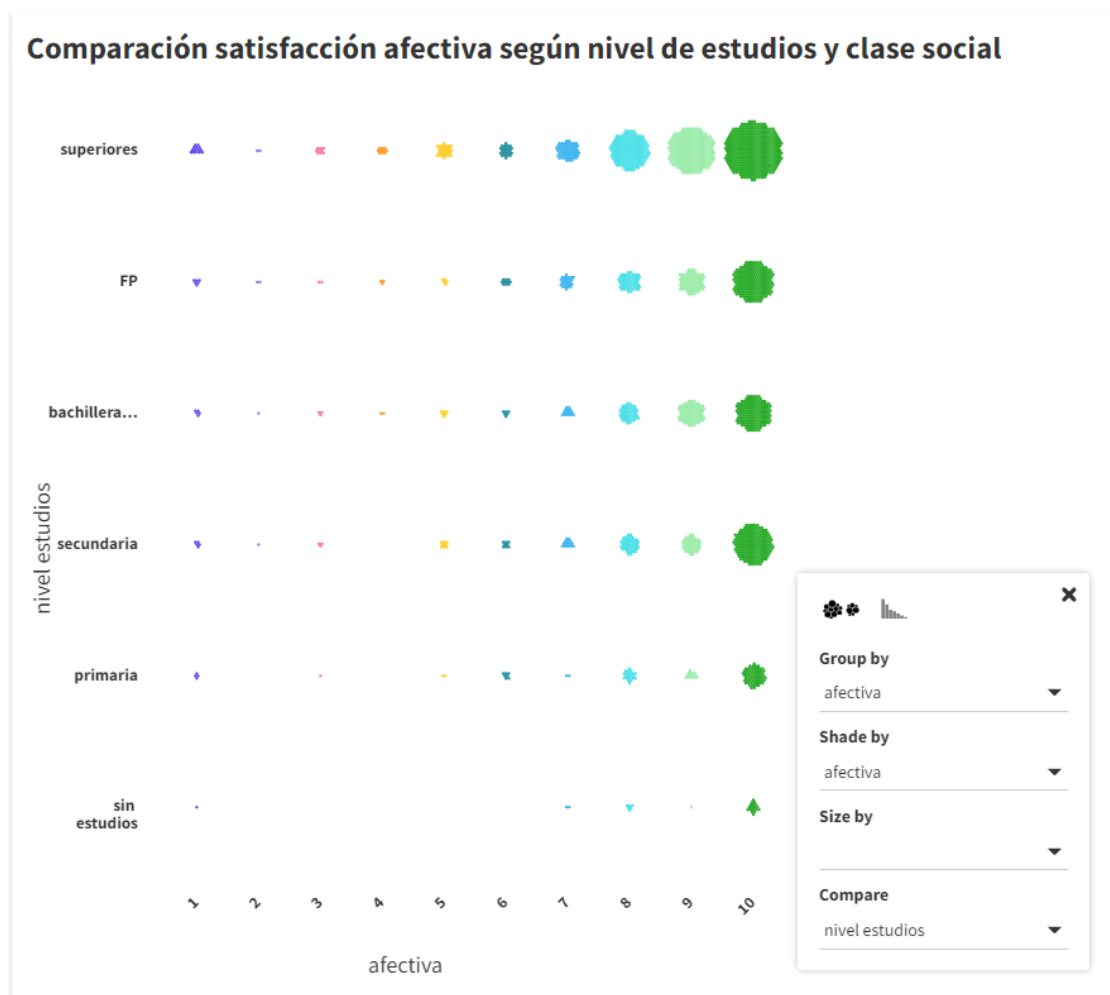
En esta ocasión vemos que los círculos están más dispersos, así como sus colores. Sin embargo, aunque de manera menos evidente, podemos llegar a distinguir una breve tendencia similar a la obtenida anteriormente, pues hay más cantidad de círculos oscuros cuánto mayor es el valor de la satisfacción social y afectiva.

2.3.- Satisfacción afectiva según clase social y nivel de estudios

Para contestar esta pregunta, se ha optado por hacer una visualización interactiva de tipo survey en Flourish, la cual consiste en un 'dot visualisation' que nos permite comprobar fácilmente la distribución de las respuestas en una pregunta de tipo encuesta, así como comparar el resultado de la cuestión con otros atributos. Podemos consultarlo en el siguiente enlace:

<https://public.flourish.studio/visualisation/12524247/>

Agrupando y coloreando los puntos según la puntuación de cada persona sobre su satisfacción afectiva (del 1 al 10 en orden ascendente), podemos ver la distribución de dicha satisfacción según el nivel de estudios de cada individuo.



Observamos que, proporcionalmente, la distribución en la puntuación obtenida de la satisfacción afectiva es similar en todos los niveles de educación, por lo que no podemos determinar que haya una tendencia clara.

Comprobaremos ahora la distribución de la satisfacción afectiva de las personas según su clase social, siguiendo los mismos criterios de color y de agrupación que anteriormente.



Nuevamente, observamos que la distribución en la puntuación de satisfacción afectiva es similar independientemente de la clase social a la que pertenecen los individuos, por lo que no podemos afirmar que haya una tendencia y tendremos que suponer que son variables independientes.

La elección de la paleta de colores es la predeterminada por el software, pues nos distingue cada valor disponible con un color ligeramente parecido al que tiene al lado, pero lo suficientemente distinto para notar que hay una clara diferencia entre las distintas agrupaciones.