

STORYTELLING WITH DATA: CH 2

JOSE ALFREDO GARCIA GARCIA

OBJETIVO



El objetivo principal para una presentación es poder transmitir a la audiencia de forma efectiva una idea o resultados de la investigación.

¿Cuál es la forma más efectiva de lograrlo? ¿Cómo podemos hacerlo? ¿Qué debemos evitar?

A continuación responderemos cada una de las preguntas.



¿Cuál es la mejor gráfica?

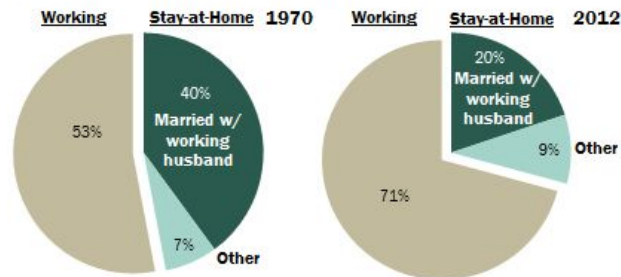
Si lo que buscamos es informar una cifra o queremos compartir información de forma directa la mejor gráfica será un simple número.

20%

20% de niños tienen a una mamá “tradicional” en 2012 comparado con el 40% en 1970.

Stay-at-Home and Working Mothers, 1970 and 2012

% of mothers with child(ren) younger than 18 who are ...



Note: Based on mothers ages 18-69 with own child(ren) younger than 18 in the household. Mothers are categorized based on employment status in 1970 and 2012. "Other" stay-at-home mothers are those who are single, cohabiting, or married with a non-working or absent husband.

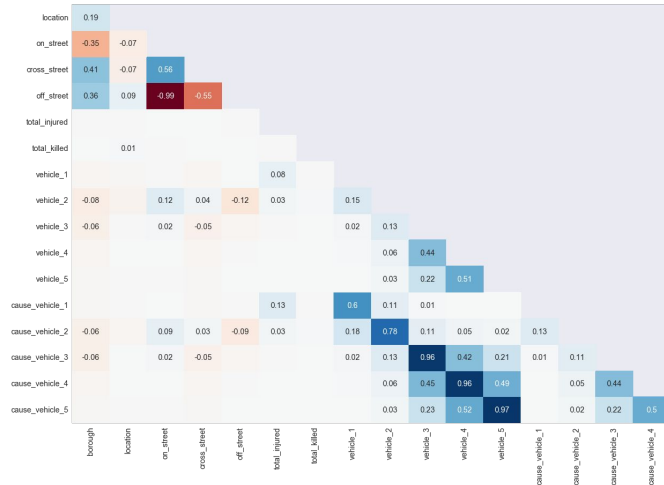
Source: Pew Research Center analysis of March Current Population Surveys Integrated Public Use Microdata Series (IPUMS-CPS), 1971 and 2013

PEW RESEARCH CENTER



Tablas y heatmaps

Una forma efectiva de comunicar múltiples tipos de datos o medidas, es fácil de mostrar a través de una tabla.



Measures	1	2	3	4	5	6	7	8	9	10	11	12
1. TBDI index	—											
2. Obsessiveness	.79	—										
3. Hostility	.72	.56	—									
4. Sensitivity	.86	.66	.59	—								
5. Depression	.91	.66	.53	.74	—							
6. Anxiety	.83	.59	.55	.66	.73	—						
7. Paranoid ideation	.68	.42	.42	.55	.54	.49	—					
8. MSPSS total	-.42	-.27	-.30	-.31	-.42	-.32	-.31	—				
9. Family	-.34	-.21	-.25	-.26	-.32	-.25	-.26	.80	—			
10. Friends	-.33	-.18	-.24	-.23	-.34	-.25	-.27	.78	.35	—		
11. Significant others	-.36	-.25	-.24	-.27	-.37	-.28	-.23	.86	.59	.54	—	
12. R-UCLA-LS	.55	.34	.35	.49	.50	.46	.56	-.48	-.30	-.51	-.36	—
Mean score	1.16	1.13	1.20	.87	1.21	1.48	1.77	63.0	21.4	19.3	22.3	43.2
SD	.70	.83	.87	.79	.89	.90	.65	16.6	7.0	6.9	6.4	11.3

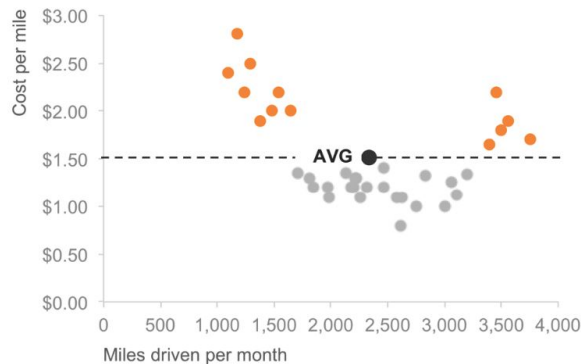
NOTE. All correlations are based on a sample size of 386 (correlation's in excess of .14 are significant at $P < .05$).
Abbreviations: TBDI, Talbieh Brief Distress Inventory; MSPSS, Multidimensional Scale of Perceived Social Support (Family, Friend and Significant other subscales); R-UCLA-LS, Revised UCLA Loneliness Scale.

SCATTERPLOTS & LINE PLOTS

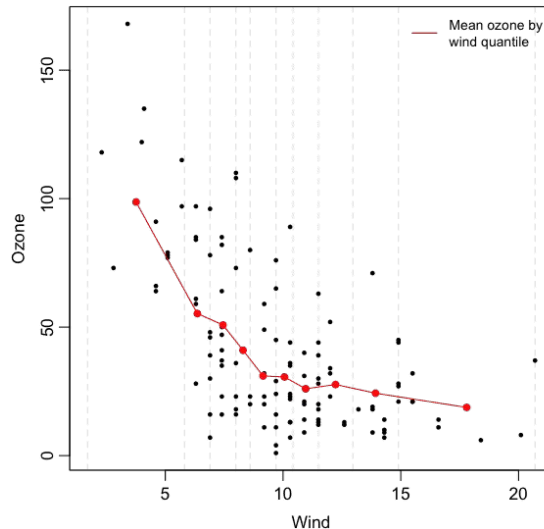
Una gráfica bien diseñada es mejor que una tabla bien diseñada.

- Scatterplots: nos ayudarán para mostrar una relación entre dos variables.
- line plots: funcionan bien para mostrar tendencias a través del tiempo, como evolución de ventas o temperatura.

Cost per mile by miles driven



Si lo que buscamos es mostrar casos en los que el costo por milla es mayor a la media se podría usar un gráfico de este estilo.



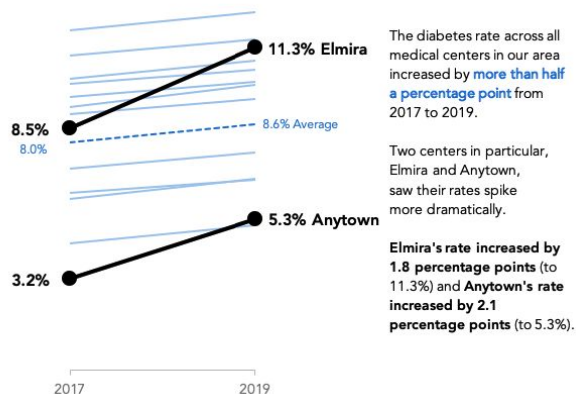


SLOPEGRAPH

Los gráficos de pendiente son mayormente usados para mostrar los cambios de valores en dos casos distintos (por ejemplo dos fechas diferentes). Son fáciles de entender ya que permiten ver rápidamente si un valor aumenta o disminuye.

Diabetes rate at all regional medical centers

PERCENTAGE OF PATIENTS WITH DIABETES



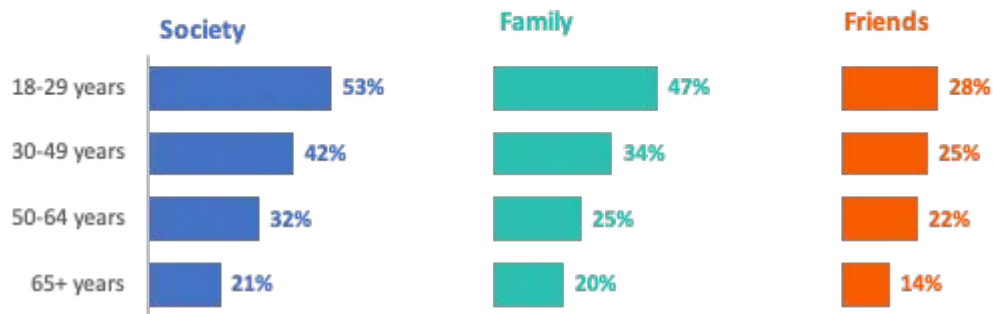


BAR PLOTS

Las gráficas de barras usualmente se evitan ya que son comunes y es un error. Las gráficas de barras son fáciles de leer para la audiencia, podemos identificar de manera sencilla los incrementos y poder realizar una comparativa.

Over half of young adult singles between ages 18 to 29 years feel that society is pressuring them to be in a relationship.

Data source: Pew Research Center, Survey of U.S. Adults, October 16-28, 2019.





LO QUE DEBEMOS EVITAR...

PIE CHART

Una de las principales dificultades al leer una gráfica de pastel es que los segmentos están representados mediante “rebanadas”, lo cual se vuelve complicado de estimar a simple vista. En otro caso, si se colocan las etiquetas de los segmentos, es complicado mostrar o ver el detalle de los porcentajes pequeños.

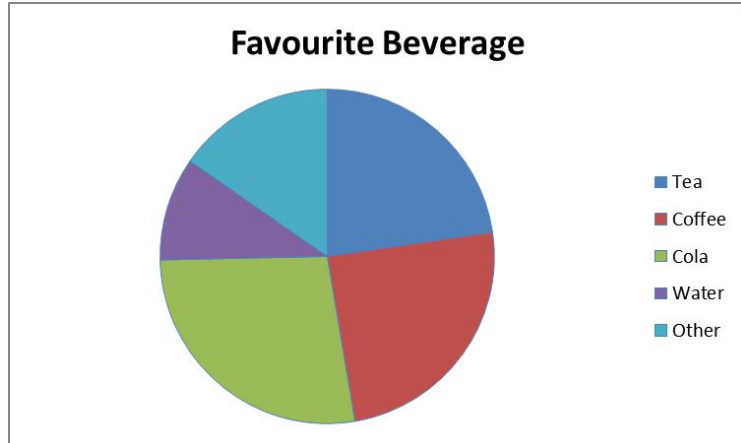
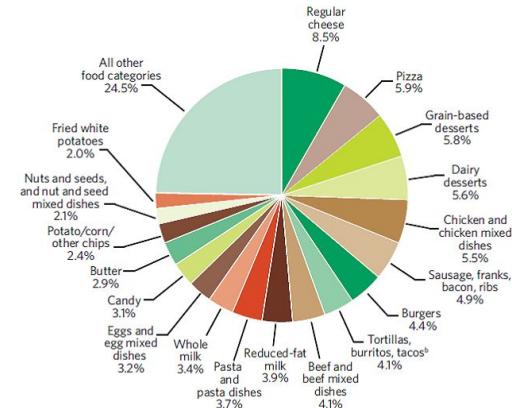


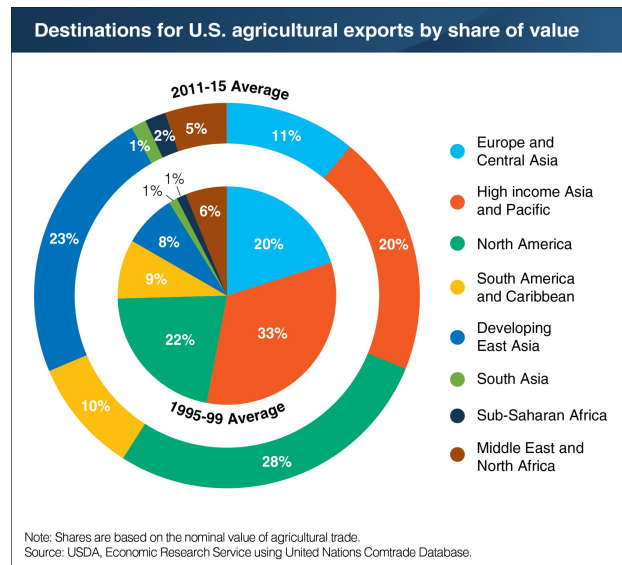
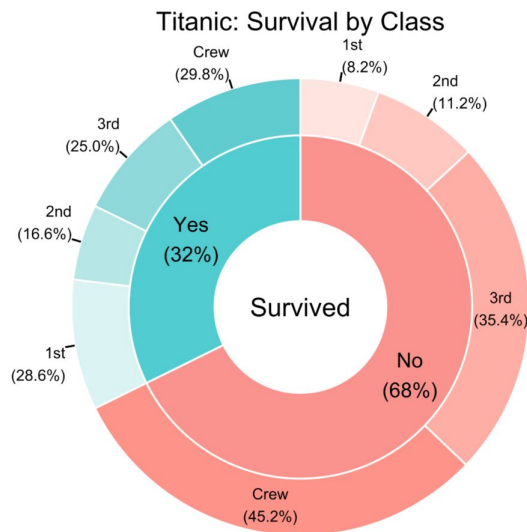
FIGURE 3-4. Sources of Saturated Fat in the Diets of the U.S. Population Ages 2 Years and Older, NHANES 2005–2006^a





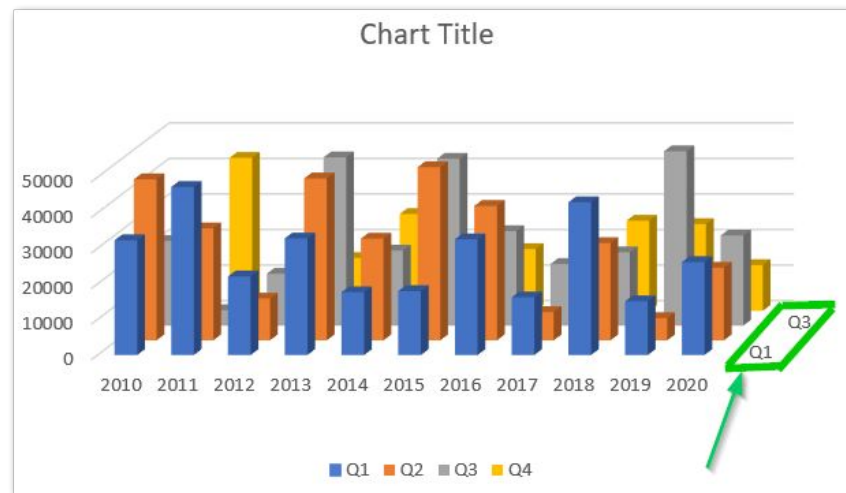
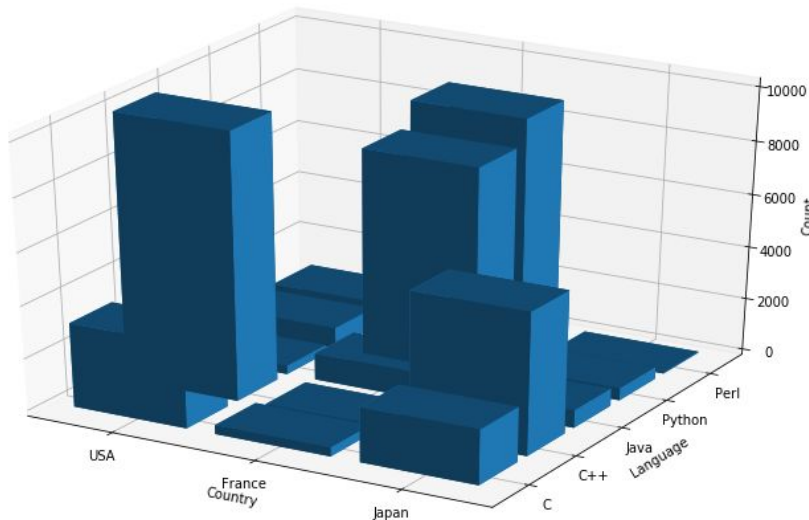
DONUT CHART.

En una grafica de pastel se pide que se comparen ángulos y áreas, en una gráfica circular se está pidiendo comparar longitudes de arcos, lo cual en sí ya era una tarea difícil ahora será peor.



EVITA PLOTS EN 3D

Una regla de oro en data visualization es nunca usar gráficas en 3D, solo en verdad si es necesario. Se a visto que las gráficas en 3D pierden sentido al no poder interpretar los resultados e incluso imposible.





CONCLUSIONES FINALES

En muchos casos no existe un solo gráfico para mostrar, por lo cual lo más importante es mostrar lo que necesitamos de una manera clara y lo más sencilla posible.

¿Qué necesitas que tu audiencia conozca? Después puedes elegir la mejor gráfica para responder.

Si te estás preguntando ¿cuál es la mejor gráfica para mi situación? La respuesta es aquella que sea fácil de leer para tu audiencia.