

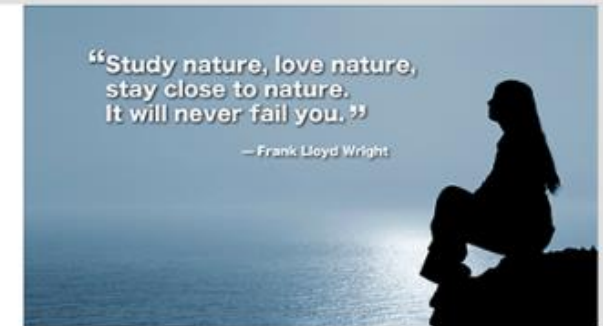
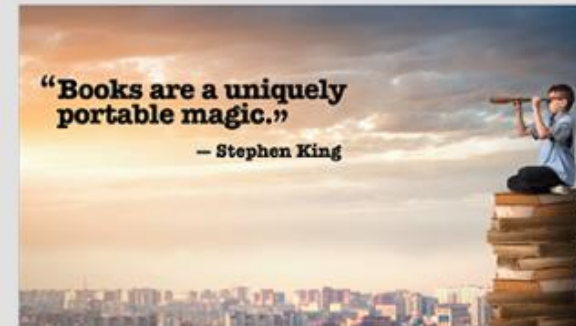
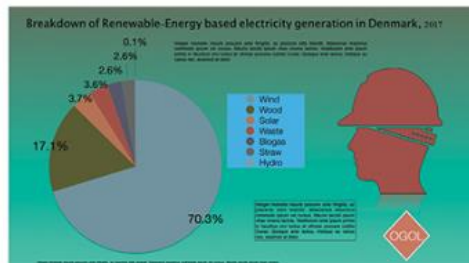
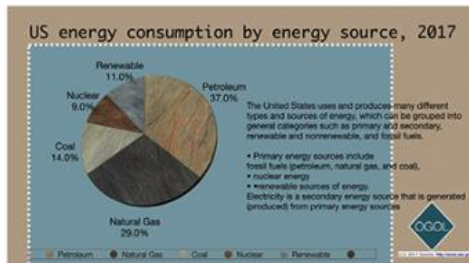
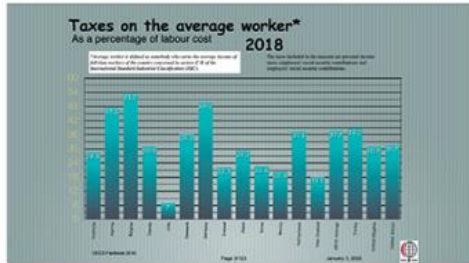
CIENCIA DE DATOS

GRADO EN BIOMEDICINA

DR. ARIEL CARIAGA-MARTÍNEZ

¿Qué diferencia observan?

BEFORE ▾



Reynolds – Presentación Zen

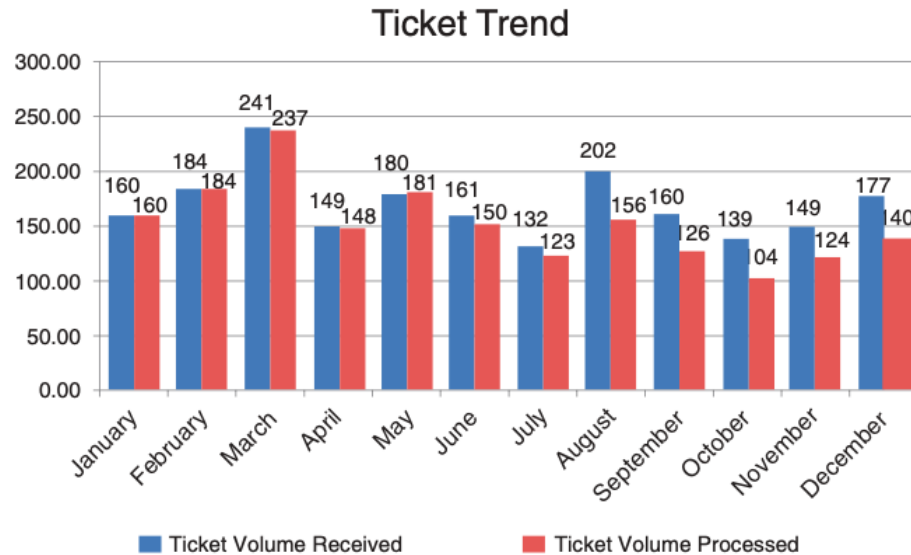
A Data Story...

Whoever your audience is, the story you are trying to tell (or the answer you are trying to share) is known to you at the outset, and therefore you can design to specifically accommodate and highlight that story. In other words, you'll need to make certain **editorial decisions** about what information stays in, and which is distracting or irrelevant and should come out. This is a process of selecting focused data that will support the story you are trying to tell.

“Who? What? Where? When? Why? How?”

**TENEMOS QUE VER DESDE EL PUNTO DE VISTA DE
NUESTRO INTERLOCUTOR**

No somos buenos con el "data storytelling"



Please approve the hire of 2 FTEs

to backfill those who quit in the past year

Ticket volume over time



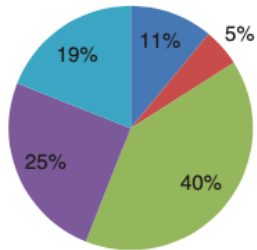
Una evolución en el tiempo
¿buena?

Data source: XYZ Dashboard, as of 12/31/2014 | A detailed analysis on tickets processed per person and time to resolve issues was undertaken to inform this request and can be provided if needed.

Survey Results

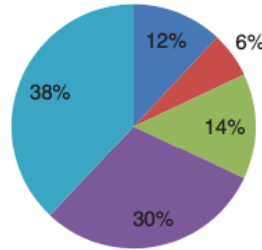
PRE: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited



POST: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited

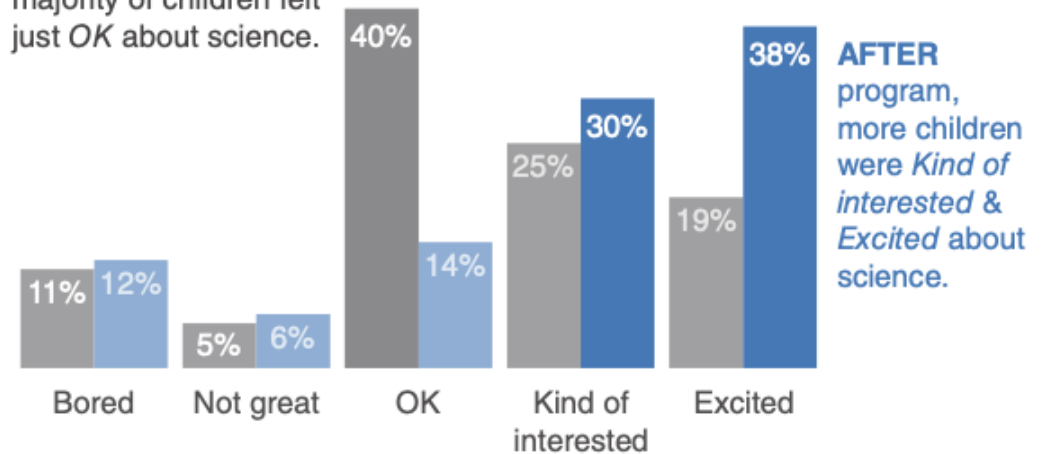


1. Understand the context
2. Choose an appropriate visual display
3. Eliminate clutter
4. Focus attention where you want it
5. Think like a designer
6. Tell a story

Pilot program was a success

How do you feel about science?

BEFORE program, the majority of children felt just *OK* about science.

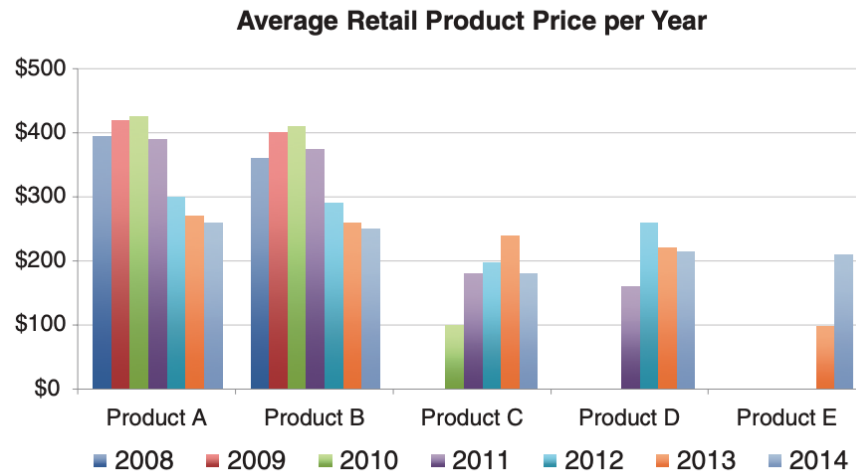


Based on survey of 100 students conducted before and after pilot program (100% response rate on both surveys).

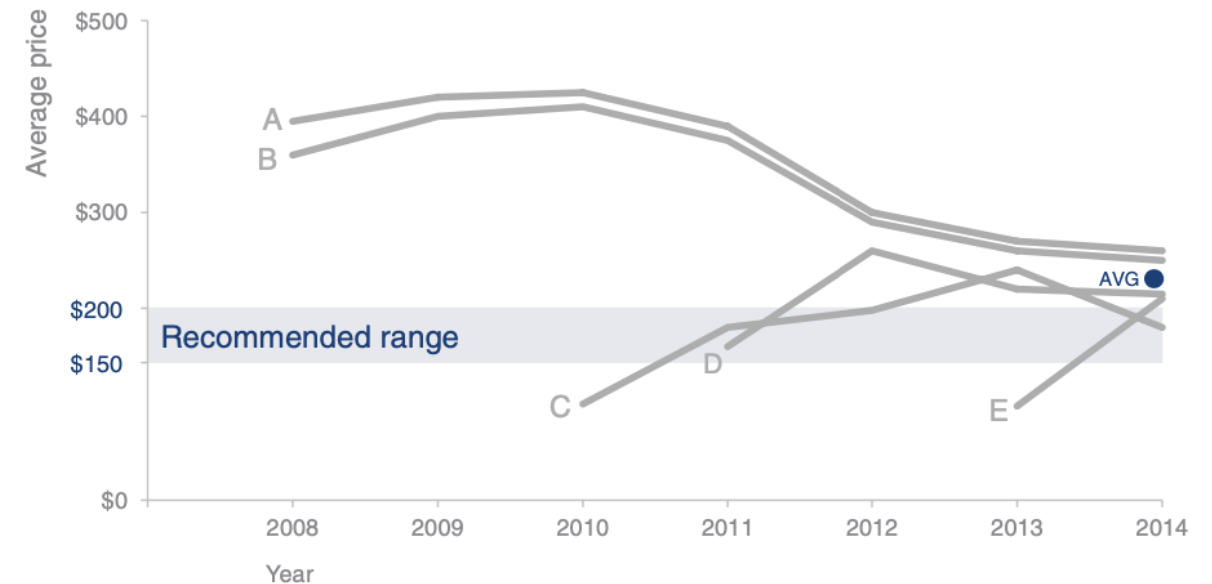


Comprender el contexto: ¿Qué significa lo que veo?

To be competitive, we recommend introducing our product *below the \$223 average price point in the \$150–\$200 range*



Retail price over time by product



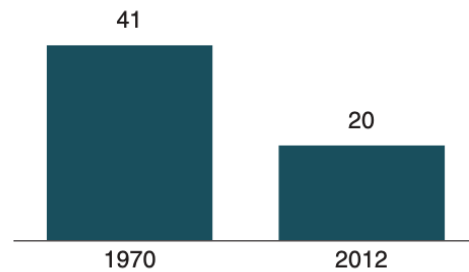
1) Adaptarse al contexto

- ¿Qué información es relevante?
- ¿Quién es la audiencia y qué sé sobre ellos?
- ¿Qué sesgos tiene mi audiencia y qué resistencias hay?
- ¿Hay datos que soporten las conclusiones? ¿Son datos “conocidos” por nuestra audiencia?
- ¿Qué riesgos hay? Puntos débiles que afectan nuestra historia.
- ¿Podemos resumir el mensaje en una línea?

2) Seleccionar un sistema de visualización efectivo

Children with a "Traditional" Stay-at-Home Mother

% of children with a married stay-at-home mother with a working husband



Note: Based on children younger than 18. Their mothers are categorized based on employment status in 1970 and 2012.

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

Adapted from PEW RESEARCH CENTER

20%

of children had a
traditional stay-at-home mom
in 2012, compared to 41% in 1970

USO ADECUADO DEL COLOR

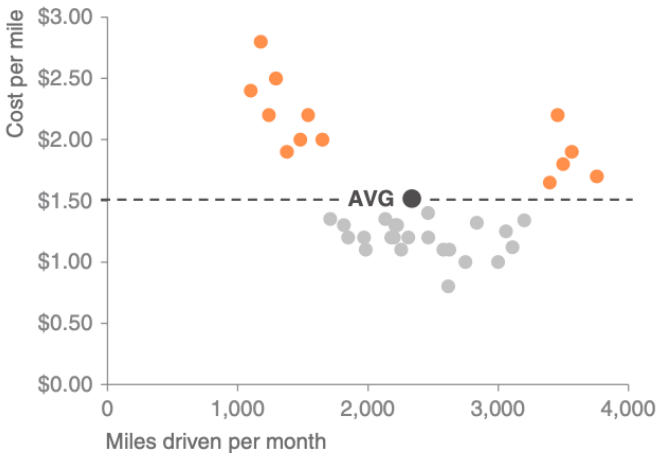
USO ADECUADO DEL COLOR I

Employee feedback over time



USO ADECUADO DEL COLOR II

Cost per mile by miles driven



TABLAS VS HEATMAPS

Table

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Heatmap

LOW-HIGH			
	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

INFOGRAFÍA VS VISUALIZACIÓN DE DATOS

EXPLORACIÓN VS EXPLICACIÓN

INFOGRAFÍA: PRIMA EL ASPECTO ESTÉTICO.

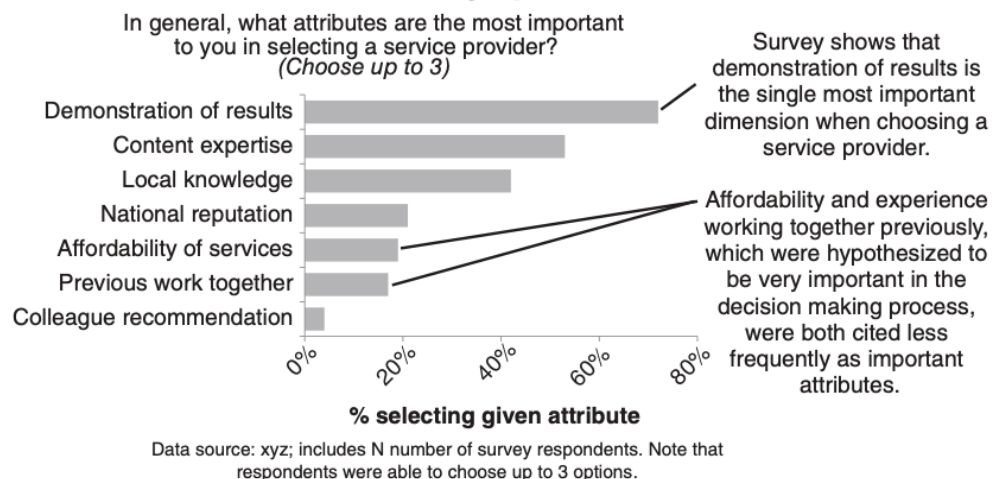
VISUALIZACIÓN DE DATOS: GENERADOS DE FORMA “ALGORITMICA”, GENERALMENTE NO ESTÁ “DECORADO”. CASI SIEMPRE MÁS FÁCIL DE SER “RE-GENERADOS”, MODIFICADOS Y ENRIQUECIDOS CON DETALLES.

GRÁFICOS EXPLORATORIOS: PARA RECORRER Y COMPRENDER EL DATASET
VS

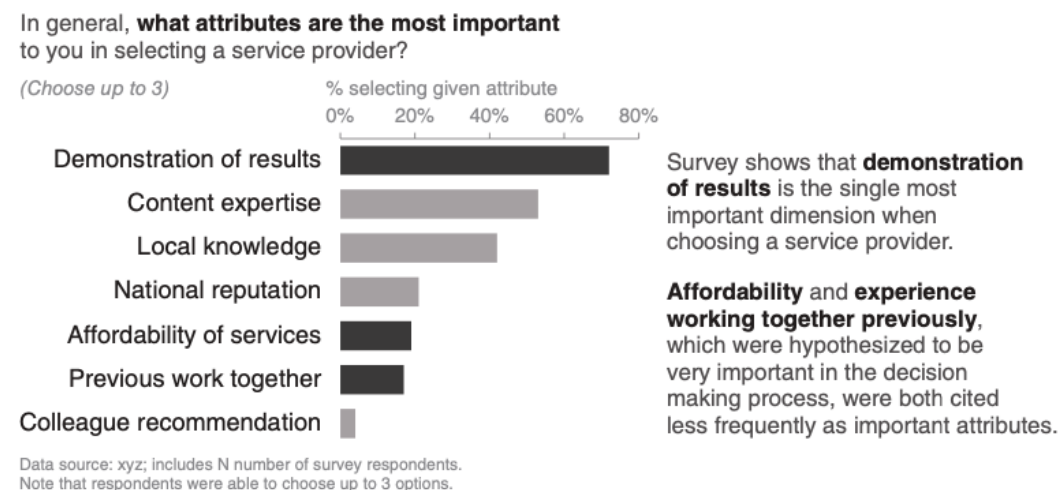
GRÁFICOS EXPLICATIVOS: PARA “TRANSPORTAR” EL CONOCIMIENTO DE NUESTRO TRABAJO A OTROS (CONCLUSIONES)

Carga (cognitiva): SEÑAL/RUIDO

Demonstrating effectiveness is most important consideration when selecting a provider

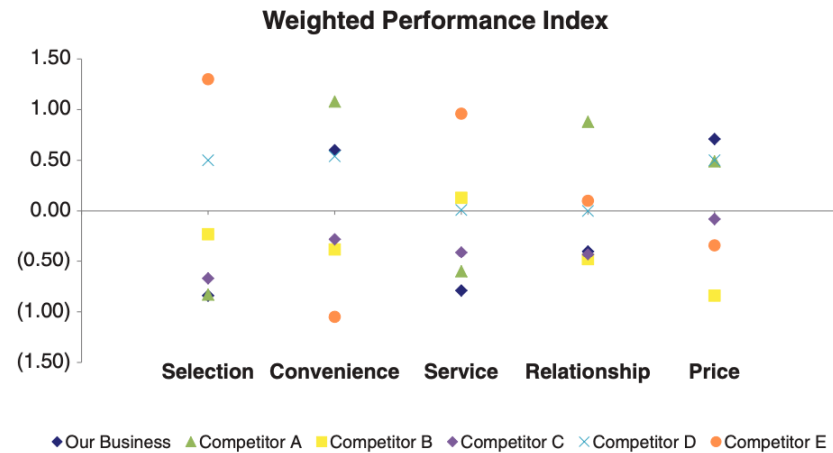


Demonstrating effectiveness is most important consideration when selecting a provider



Problema: carencia de orden/jerarquía visual

El color también puede ser desorden.

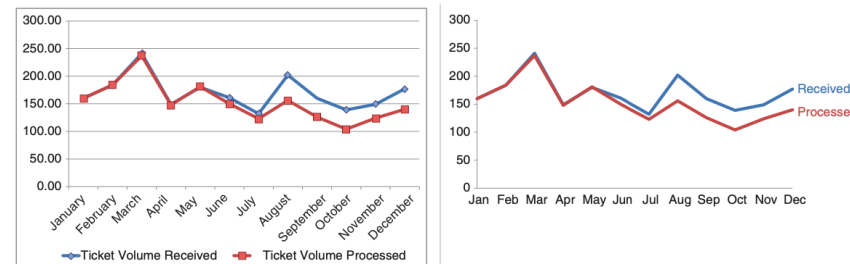
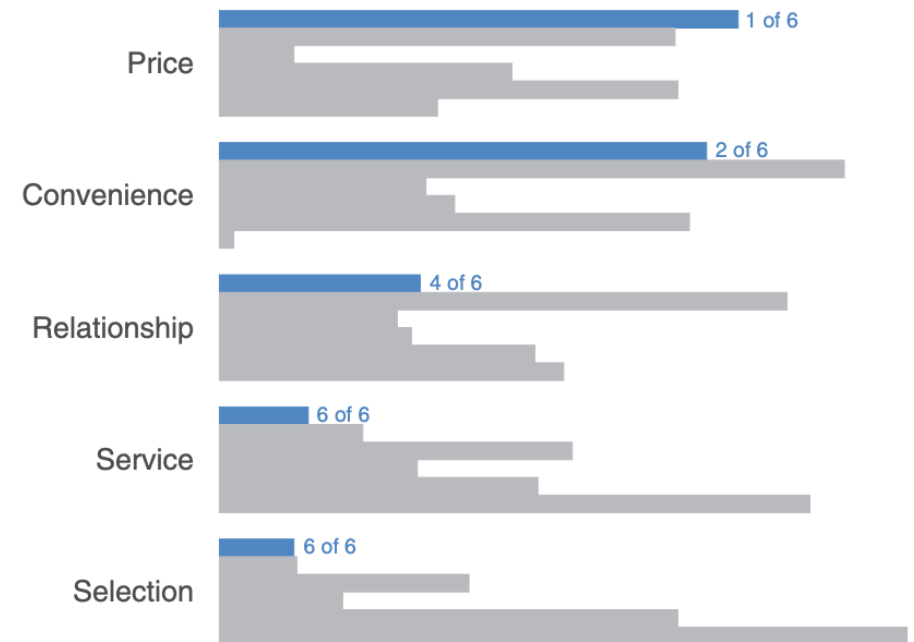


Performance overview

Our business

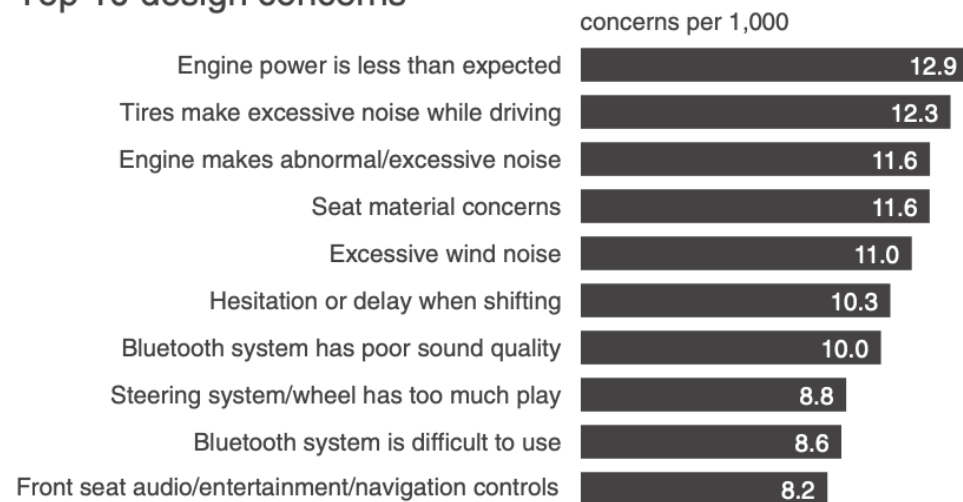
- Competitor A
- Competitor B
- Competitor C
- Competitor D
- Competitor E

Weighted performance index | relative rank

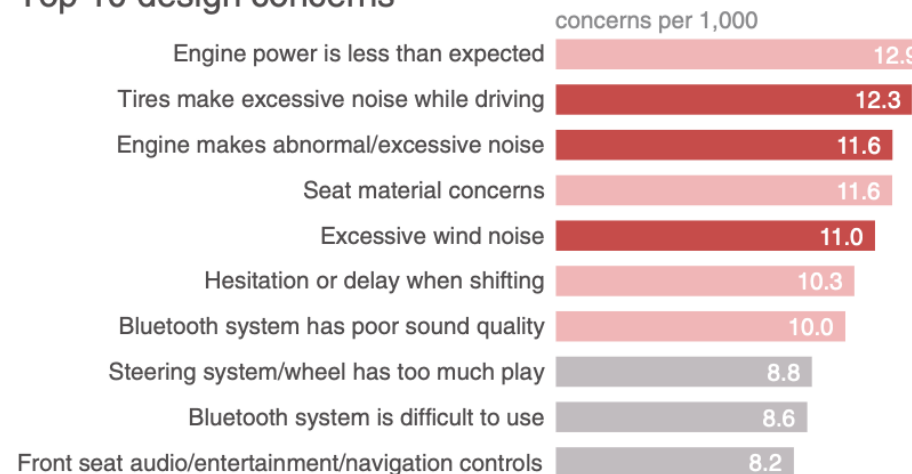


Carga (cognitiva): SEÑAL/RUIDO

Top 10 design concerns



Top 10 design concerns



Comments indicate that **noisy tire issues** are most apparent **in the rain**.

Complaints about **engine noise** commonly cited **after the car had not been driven for a while**.

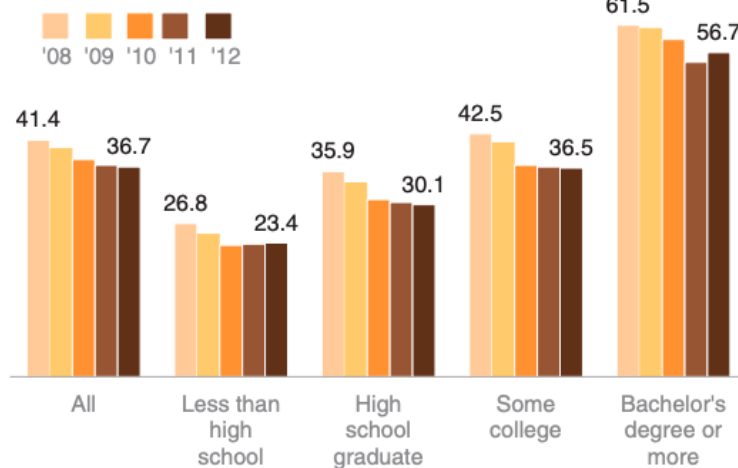
Excessive **wind noise** is noted primarily in **freeway driving at high speeds**.

Enfocar la atención de la audiencia

Carga (cognitiva): SEÑAL/RUIDO

New Marriage Rate by Education

Number of newly married adults per 1,000 marriage eligible adults



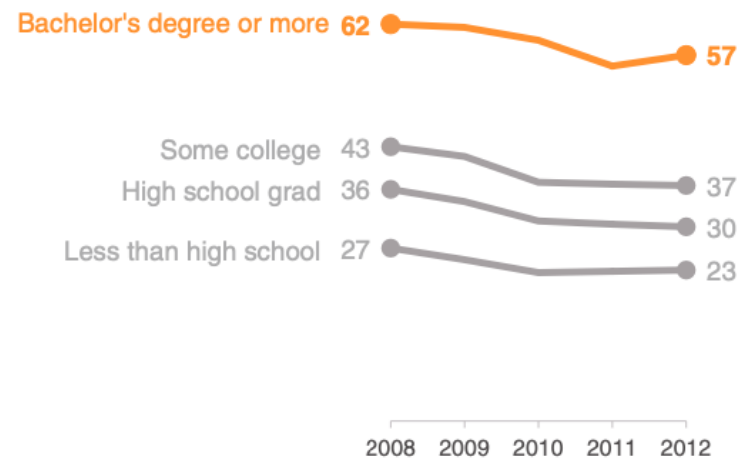
Note: Marriage eligible includes the newly married plus those widowed, divorced, or never married at interview.

Source: U.S. Census

Adapted from PEW RESEARCH CENTER

New marriage rate by education

Number of newly married adults per 1,000 marriage eligible adults



Note: Marriage eligible includes the newly married plus those widowed, divorced, or never married at interview.

Source: U.S. Census

Adapted from PEW RESEARCH CENTER

Enfocar la atención de la audiencia

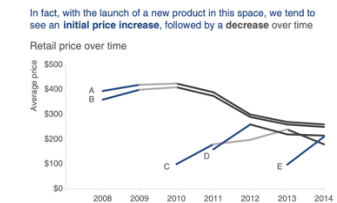
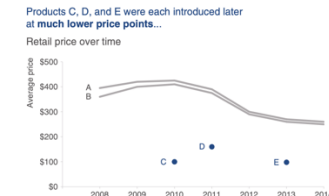
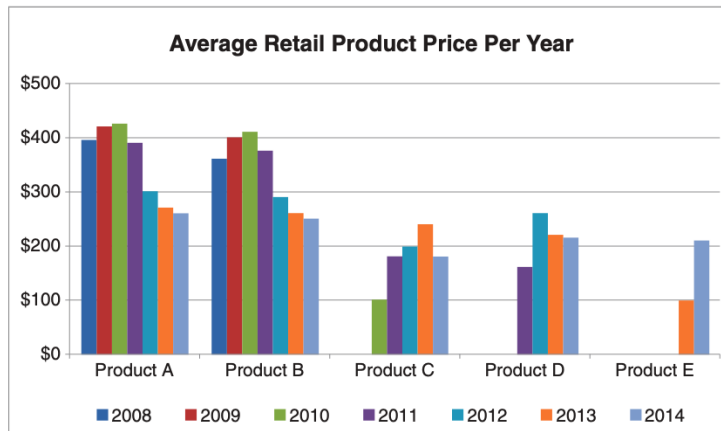
In the next 5 minutes...

OUR GOAL:

- 1 Understand **how prices have changed over time** in the competitive landscape.
- 2 Use this knowledge to **inform the pricing of our product**.

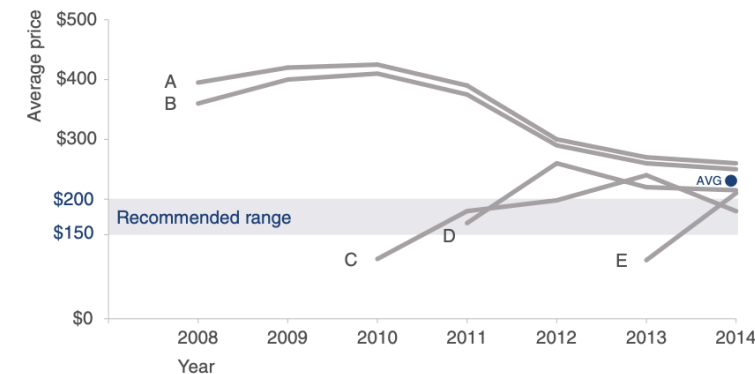
We will end with a **specific recommendation**.

Price has declined for all products on the market since the launch of Product C in 2010



To be competitive, we recommend introducing our product *below the \$223 average price point in the \$150–\$200 range*

Retail price over time



Claves

- **10': Storytelling.**
- **¿Qué podemos encontrar con respecto a...?**
- **Comunicación.**
- **Contar una historia interesante.**
- **Obtener conclusiones relevantes.**
- **Información a conocimiento.**

GRACIAS

DR. ARIEL CARIAGA-MARTINEZ

CIENCIA DE DATOS

ACARIMAR@UAX.ES

TRABAJO

GRADO EN BIOMEDICINA

DR. ARIEL CARIAGA-MARTÍNEZ

Una mejora constante...

- Formar equipos (2-3 personas).
- Escoger un *dataset* o continuar con el del trabajo anterior.
- RECOMENDACIÓN: <https://archive.ics.uci.edu>
- A partir del EDA generar al menos 4 gráficos **explicativos** a partir de los gráficos exploratorios, CONTANDO UNA HISTORIA.
- Generar al menos los bocetos de los gráficos explicativos comentando las mejoras propuestas. SE PUEDE USAR EXCEL/POWERPOINT
 - RETO: ESCRIBIR EL CÓDIGO EN EL PAQUETE ggplot2/Seaborn
- NO CHATGPT: NO SE TRATA DE HACER POR HACER SINO MEJORAR EL PODER EXPLICATIVO.
- PRESENTAR LOS RESULTADOS DE UNA FORMA VISUAL/INTERACTIVA (MÁXIMO 2 FOLIOS SI ES “ESCRITO”: MEDIO FOLIO POR GRÁFICA). **ANTES/DESPUÉS**
- AGREGAR UNA PEQUEÑA JUSTIFICACIÓN DE LA APLICACIÓN DE PRINCIPIOS DE STORYTELLING.
- Generar una versión PARA CEOs y una versión para matemáticos/científicos de datos.



GUÍA DE PREGUNTAS PARA STORYTELLING CON DATOS

1. ¿Quién es nuestra audiencia objetivo? La respuesta a esta pregunta influirá tanto en la longitud como en el contenido de nuestra historia de datos.
2. ¿Qué fuente de datos estamos examinando y qué concluimos de ella? Dado que estamos describiendo los resultados del análisis de datos, es importante ser conciso, pero no ambiguo respecto a cuáles son nuestras conclusiones y los datos en los que se basan;
3. ¿De dónde provienen estos datos? Dado que "los mismos" datos pueden provenir de varias fuentes diferentes, es importante al menos proporcionar una referencia a la fuente de los datos en los que se basa nuestra historia;
4. ¿Cuándo se generaron estos datos? Incluso en casos donde el tiempo no entra explícitamente en nuestro análisis, saber cuándo se recopilaron los valores de los datos puede ser importante (por ejemplo, ¿estos datos de ventas nacionales se recopilaron antes o después de la Gran Recesión?);
5. ¿Por qué estamos analizando este conjunto de datos? Junto con las preguntas sobre qué fuente de datos estamos analizando y qué concluimos de ella, esta es probablemente la pregunta más importante que debe abordarse en cualquier historia de datos;
6. ¿Cómo llegamos a nuestras conclusiones? Las historias de datos para audiencias no técnicas no deben entrar en gran detalle técnico, pero se debe proporcionar suficiente detalle—o citar referencias a fuentes de estos detalles—para dejar al lector cómodo con que se realizó un análisis sistemático.