

## 1. Big Data, Ciencia de Datos y Visualización

### 1.1 Relación e Importancia

La visualización de datos actúa como el puente cognitivo entre el procesamiento masivo y la comprensión humana:

- **Big Data:** Permite gestionar las "V" del Big Data (Volumen, Variedad, Velocidad) al resumir conjuntos masivos de datos en patrones visuales que el ojo humano detecta en milisegundos.
- **Análisis de Datos:** Es la herramienta principal para el Análisis Exploratorio de Datos (EDA), permitiendo validar hipótesis y encontrar tendencias antes de aplicar modelos complejos.
- **Inteligencia de Datos (BI):** Facilita la toma de decisiones al transformar métricas operativas en indicadores visuales que cuentan una historia de negocio clara.
- **Ciencia de Datos:** Es fundamental en la comunicación de resultados complejos; sin una visualización efectiva, los modelos predictivos no son accionables para las partes interesadas.

### 1.2 Niveles de Abstracción y Clasificación de Datos

- **Niveles de Abstracción (Pirámide DIKW):**
  1. **Dato:** El valor atómico sin procesar (ej. "200 minutos").
  2. **Información:** El dato con contexto (ej. "El tiempo de uso de pantallas en 2022").
  3. **Conocimiento:** La integración de información para identificar patrones (ej. "El uso de pantallas ha desplazado el tiempo social").
  4. **Sabiduría:** El uso del conocimiento para tomar decisiones éticas o estratégicas (ej. "Fomentar políticas de desconexión digital").
- **Clasificación de Datos en el Mundo Real:** Los datos suelen encontrarse de forma **desestructurada** (texto, imágenes).
- **Tipos de Datos (según Munzner):**
  - **Items y Atributos:** Los elementos y sus propiedades.
  - **Links:** Relaciones entre items (comunes en grafos).
  - **Positions:** Datos espaciales o geográficos



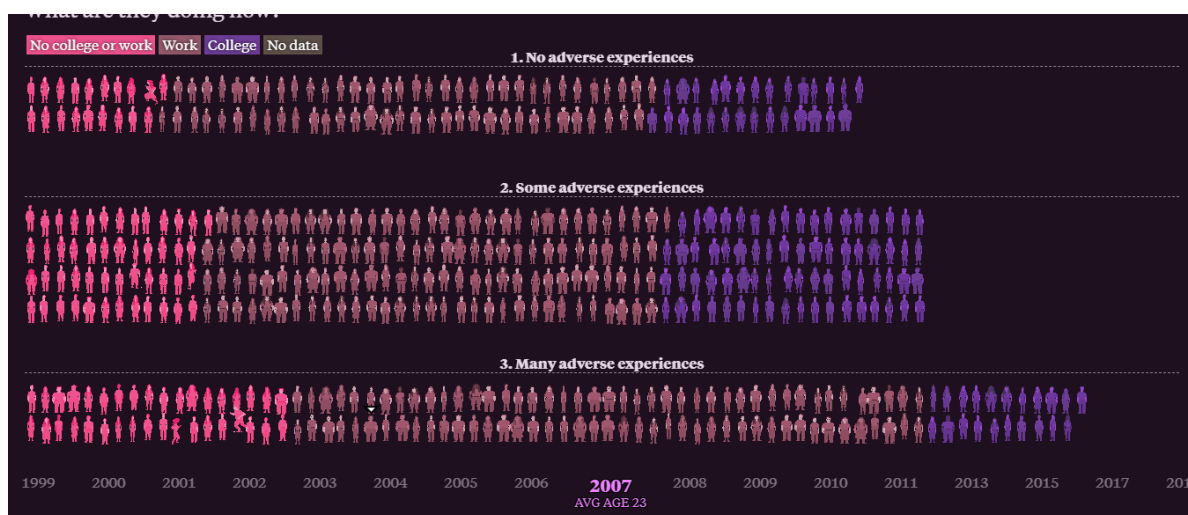
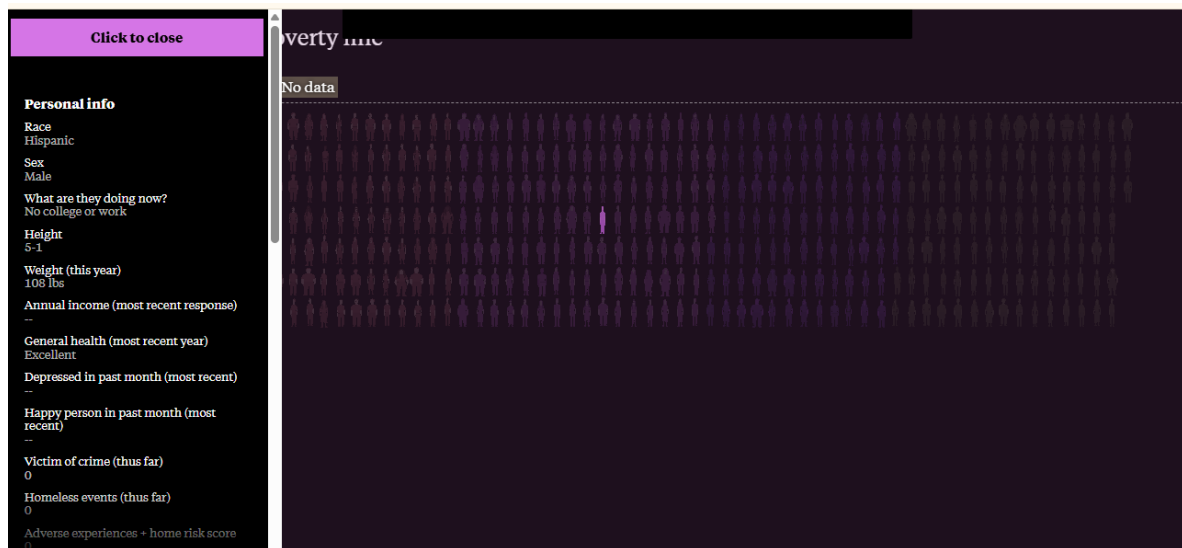
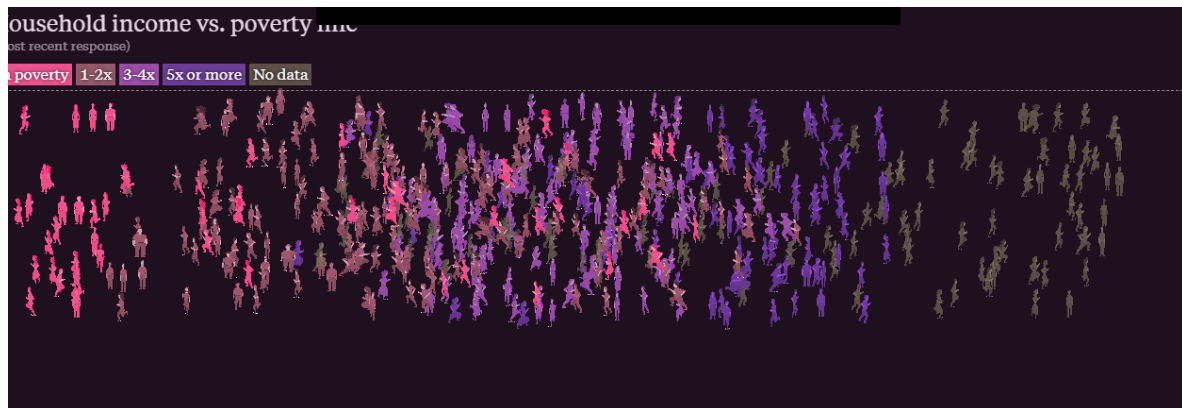
## 2. Análisis de Visualización:

### "How Teenagers Spend Their Time"

Enlace: [pudding.cool/2024/03/teenagers/](https://pudding.cool/2024/03/teenagers/)

En esta visualización tenemos la oportunidad de ver como va cambiando la vida de los adolescentes a lo largo de los años, esta es una visualización narrativa, mediante scroll:





## 2.1 WHAT

- **Actividades (Screen Time, Sleep, etc.):** Es un atributo **Categorical** (categórico/nominal). No tiene un orden intrínseco, solo distingue tipos de tareas.

- **Año (2003 - 2022):** Es un atributo **Ordered -> Ordinal**. Los años tienen un orden lógico y secuencial en el tiempo.
- **Tiempo dedicado (Minutos):** Es un atributo **Quantitative**. Representa una magnitud numérica exacta.
- **Dataset Type:** Se clasifica como una **Table** (tabla) con atributos temporales y categóricos.

## 2.2 WHY

Un **modismo (idiom)** es la combinación de codificación visual e interacción. Se analizaron dos:

### 1. Modismo: Gráfico de Áreas Apiladas Dinámico

Aquí compramos o tenemos un recuento de tendencias y distribuciones, **Compare Trends or Summarize Distribution**, así:

- **Action : Compare y Summarize.** El usuario busca comparar cómo cambia la distribución del tiempo entre décadas.
- **Target: Trends y Distribution.** El objetivo es identificar la tendencia al alza del tiempo digital frente a otras actividades.

### 2. Modismo: Scrollytelling (Interacción de Navegación)

Aquí vemos en esta visualización Presentar Características o **Present Features**, así:

- **Action : Consume -> Present.** El autor guía al usuario para presentar una historia basada en datos.
- **Target : Features.** Ayuda a localizar eventos o hallazgos específicos en los datos (como el impacto de la pandemia).

## 2.3 Conclusiones y Evaluación

- Es una visualización de alta calidad porque aplica correctamente el **Mapping de Color (Hue)** para atributos categóricos y la **Posición** para el tiempo, que son los canales más efectivos según la jerarquía de Munzner. La interactividad de navegación (scroll) evita la sobrecarga cognitiva y ruido visual excesivo.
- Puede considerarse limitada si el usuario busca **Locate** (localizar) valores exactos, ya que las áreas apiladas dificultan la lectura precisa de los valores intermedios sin una herramienta de **Select** (selección) o tooltip que muestre el dato exacto al pasar el cursor