

Concurrencia a Museos de Argentina

Relación con otros consumos culturales

Nazareno Magallanes Javier Spina Lautaro Terreno

Escuela de Ciencia y Tecnología

Universidad Nacional de San Martín

Noviembre 2022

Presentación del *dataset*

- Se trabajó con el *dataset* asociado a la **Encuesta Nacional de Consumos Culturales** realizada en 2017. Son datos abiertos del Ministerio de Cultura de la Nación, disponible en su página web.

Presentación del *dataset*

- Se trabajó con el *dataset* asociado a la **Encuesta Nacional de Consumos Culturales** realizada en 2017. Son datos abiertos del Ministerio de Cultura de la Nación, disponible en su página web.
- Además del archivo csv, se contó con el cuestionario aplicado y el Informe General realizado por el mismo Ministerio.

Presentación del *dataset*

- Se trabajó con el *dataset* asociado a la **Encuesta Nacional de Consumos Culturales** realizada en 2017. Son datos abiertos del Ministerio de Cultura de la Nación, disponible en su página web.
- Además del archivo csv, se contó con el cuestionario aplicado y el Informe General realizado por el mismo Ministerio.
- El *dataset* consta de 2802 unidades de análisis y 450 variables. Cada una de las variables se corresponde con las preguntas del cuestionario asociado.

Presentación del *dataset*

- Se trabajó con el *dataset* asociado a la **Encuesta Nacional de Consumos Culturales** realizada en 2017. Son datos abiertos del Ministerio de Cultura de la Nación, disponible en su página web.
- Además del archivo csv, se contó con el cuestionario aplicado y el Informe General realizado por el mismo Ministerio.
- El *dataset* consta de 2802 unidades de análisis y 450 variables. Cada una de las variables se corresponde con las preguntas del cuestionario asociado.
- Cada unidad de análisis es una persona encuestada.

Un primer vistazo

- Al comenzar a trabajar con el *dataset*, se encontraron diversas tareas para hacer antes de hacer una manipulación concreta de los datos. Fundamentalmente, los nombres de las variables:

```
> colnames(encuesta)
[1] "id"      "pondera_dem"  "fecha"  "region"
[5] "sexo"    "edad"        "p1"     "p1otros"
[9] "p2"      "p2_1"        "p2_1otro" "p2_2"
[13] "p3"      "p4"          "p5"     "p6horas"
```

Un primer vistazo

- Al comenzar a trabajar con el *dataset*, se encontraron diversas tareas para hacer antes de hacer una manipulación concreta de los datos. Fundamentalmente, los nombres de las variables:

```
> colnames(encuesta)
[1] "id"    "pondera_dem"  "fecha"  "region"
[5] "sexo"  "edad"  "p1"    "p1otros"
[9] "p2"    "p2_1"    "p2_1otro"  "p2_2"
[13] "p3"    "p4"    "p5"    "p6horas"
```

- La gran mayoría de las variables es de la forma p + número de pregunta del cuestionario.

- Para empezar, se transformaron los nombres de las variables considerando la pregunta asociada del cuestionario. Por ejemplo, p98 se renombró a concurre_museo porque se corresponde a la pregunta 98: ¿Concurrió a algún museo durante el último año?

- Para empezar, se transformaron los nombres de las variables considerando la pregunta asociada del cuestionario. Por ejemplo, p98 se renombró a concurre_museo porque se corresponde a la pregunta 98: ¿Concurrió a algún museo durante el último año?
- Se debieron eliminar 2 unidades de análisis que estaban mal cargadas.

Búsqueda de la problemática

- Ahora comenzaba la trayectoria para buscar una problemática a resolver mediante la Ciencia de Datos. Se comenzó usando un escenario hipotético para segmentar consumidores de una bebida energética y sugerir estrategias basadas en los consumos culturales.

Búsqueda de la problemática

- Ahora comenzaba la trayectoria para buscar una problemática a resolver mediante la Ciencia de Datos. Se comenzó usando un escenario hipotético para segmentar consumidores de una bebida energética y sugerir estrategias basadas en los consumos culturales.
- Luego, se intentó centrar el problema en buscar una diferencia de preferencias culturales entre nativos digitales y no-nativos digitales.

Búsqueda de la problemática

- Ahora comenzaba la trayectoria para buscar una problemática a resolver mediante la Ciencia de Datos. Se comenzó usando un escenario hipotético para segmentar consumidores de una bebida energética y sugerir estrategias basadas en los consumos culturales.
- Luego, se intentó centrar el problema en buscar una diferencia de preferencias culturales entre nativos digitales y no-nativos digitales.
- Mediante varias iteraciones de lectura, interpretación, pruebas y errores se vió que el mismo informe mostraba una problemática mostrada por la Encuesta: hay una tendencia negativa en la concurrencia a los museos argentinos en determinados sectores sociales.

La pregunta y su contexto

La pregunta

¿Cómo aumentar la concurrencia a los museos?

Contexto

Poner acá primeros gráficos parecidos a los de la página 23 del informe para mostrar contexto. Quizá alguno más puede aportar, hay que ver cuál. Quizá apoyarnos con el dataset extra de ubicaciones de los museos en el país, considerar si el factor de que la gente no tiene ni 1 auto en el nivel socio económico q no está yendo a museo.

Estrategia para responder la pregunta

Acá hablar de que buscamos de los que SI fueron a museo tuvieron ciertos preferencias en otros consumos culturales como música, tv, películas. Teniendo en cuenta esa segmentación, buscar a la gente que tiene esas preferencias y NO fue a museos y proponer esto como estrategia para alentar la concurrencia, invocando el dato de q 75 por ciento de la gente que fue a museo no pagó entrada. Osea sugerir una semana de los museos como la noche de los museos con entrada gratis para q la gente pueda ir.

Se pueden meter los modelos acá también, poner en Eje X la edad, en Eje Y NSEpuntaje (más bajo, menor clase social, más álto el puntaje mayor clase social) y buscar modelar ahí y ajustar por cantidad de autos, si concurrió o no a museo, por región, etc. No estoy seguro q sale de eso.