



PROYECTO FINAL DATA ANALISYS

ANALISIS SOBRE LA FELICIDAD A NIVEL MUNDIAL

INTEGRANTES:

Enzo Matías González
Egisto Barazarte González

pág. 1

INDICE

| | |
|--|----|
| Portada | 1 |
| Índice | 2 |
| Temática a Desarrollar | 3 |
| ¿Cómo se midió la felicidad? | 3 |
| ¿Cómo se podrían usar estos datos? | 3 |
| Temática de los Datos | 4 |
| Diagramas Entidad-Relación | 5 |
| Definición de Tipo de Campo | 6 |
| HAPPINESS | 6 |
| COUNTRIES | 6 |
| REGION | 7 |
| RELIGION | 7 |
| WEATHER | 7 |
| Transformación de Datos | 8 |
| Diagrama Relacional del Tablero | 9 |
| Análisis Funcional del Tablero | 9 |
| Medidas Calculadas | 13 |
| Herramientas Utilizadas y Futuras Líneas | 15 |

TEMATICA A DESARROLLAR

¿Cómo se midió la felicidad?

Las puntuaciones y las clasificaciones de felicidad utilizan datos de la Encuesta Mundial de Gallup. Los puntajes se basan en las respuestas a la pregunta principal de evaluación de la vida que se hizo en la encuesta. Esta pregunta, conocida como la escalera de Cantril, pide a los encuestados que piensen en una escalera en la que la mejor vida posible para ellos sea un 10 y la peor vida posible sea un 0 y que califiquen sus propias vidas actuales en esa escala. Las puntuaciones provienen de muestras representativas a nivel nacional para los años 2017-2019-2021 y, utilizan las ponderaciones de Gallup para que las estimaciones sean representativas. Las columnas que siguen al puntaje de felicidad se midieron para observar como contribuyen a hacer que las evaluaciones de vida sean más altas en cada país. No tienen ningún impacto en el puntaje total informado para cada país, pero sí explican por qué algunos países tienen una clasificación más alta que otros.

¿Cómo se podrían usar estos datos?

Objetivo

Para comprender cual es relación existente entre el índice de felicidad de cada país y los factores externos que afectan la vida de sus habitantes, como son: el HDI, ayudas sociales, la libertad, la percepción de corrupción, la pobreza, entre otros. Con ello, esperamos entender en qué medida estos factores afectan, de forma positiva o negativa, la vida de los individuos; así como también, identificar cuáles de ellos son claves en la existencia de las personas para que estas se puedan considerar, así mismos, seres felices.

Público objetivo

Se tiene como objetivo un grupo de personas con los conocimientos suficientes para comprender dichos indicadores para que estos puedan ser analizadas meticulosamente para luego ser tenidos en cuenta al momento de tomar decisiones.

Alcance

El informe posee un nivel de análisis de tipo diagnóstico y es desde la perspectiva argentina, además, se busca realizar el análisis desde el punto de vista de las variables Económicas, políticas y sociales de las que se hablará a lo largo del proyecto tales como el HDI, ayudas sociales, la libertad, la percepción de corrupción, la pobreza, entre otros.

TEMATICA DE LOS DATOS

Año (year): Año en el que se realizó la recolección de datos.

Países (country): Nombre del país.

Ranking de Felicidad (happiness_ranking): Posición del país en el ranking de felicidad según puntaje de felicidad obtenido.

Puntaje de felicidad (happiness_index): Puntaje según qué tan felices son los habitantes de un país.

Población (population): Población del país en el año considerado.

PBI anual (yearly_gdp): Producto Bruto Interno de un país en el año considerado.

PBI per cápita (yearly_gdp_pc): Producto Bruto Interno de un país por habitante en el año considerado.

Índice de Desarrollo Humano (hdi): Índice de desarrollo humano de un país en el año considerado, este abarca tres variables fundamentales: salud, educación y riqueza.

Apoyo social (social_support): Se usa una puntuación para medir que tan apoyado por su grupo social se siente una persona.

Libertad (freedom): Libertad de una persona para tomar decisiones personales.

Percepción de Corrupción (perceptions_of_corruption): Es un valor que puntúa el nivel de corrupción que cree una persona que hay en su país.

Pobreza (poverty): Porcentaje de la población que se encuentra bajo la línea de la pobreza según datos oficiales del país.

Desigualdad de Género (gender_inequality): Mide el grado de desigualdad de género basándose en la brecha salarial.

Índice de Precio al consumidor (cpi): Variación de los precios de un conjunto de productos que presentan los bienes y servicios adquiridos por un individuo.

Índice de Paz (peace_index): Puntaje en función de los conflictos internos que sufra un país.

Latitud (latitude): Latitud de la capital del país.

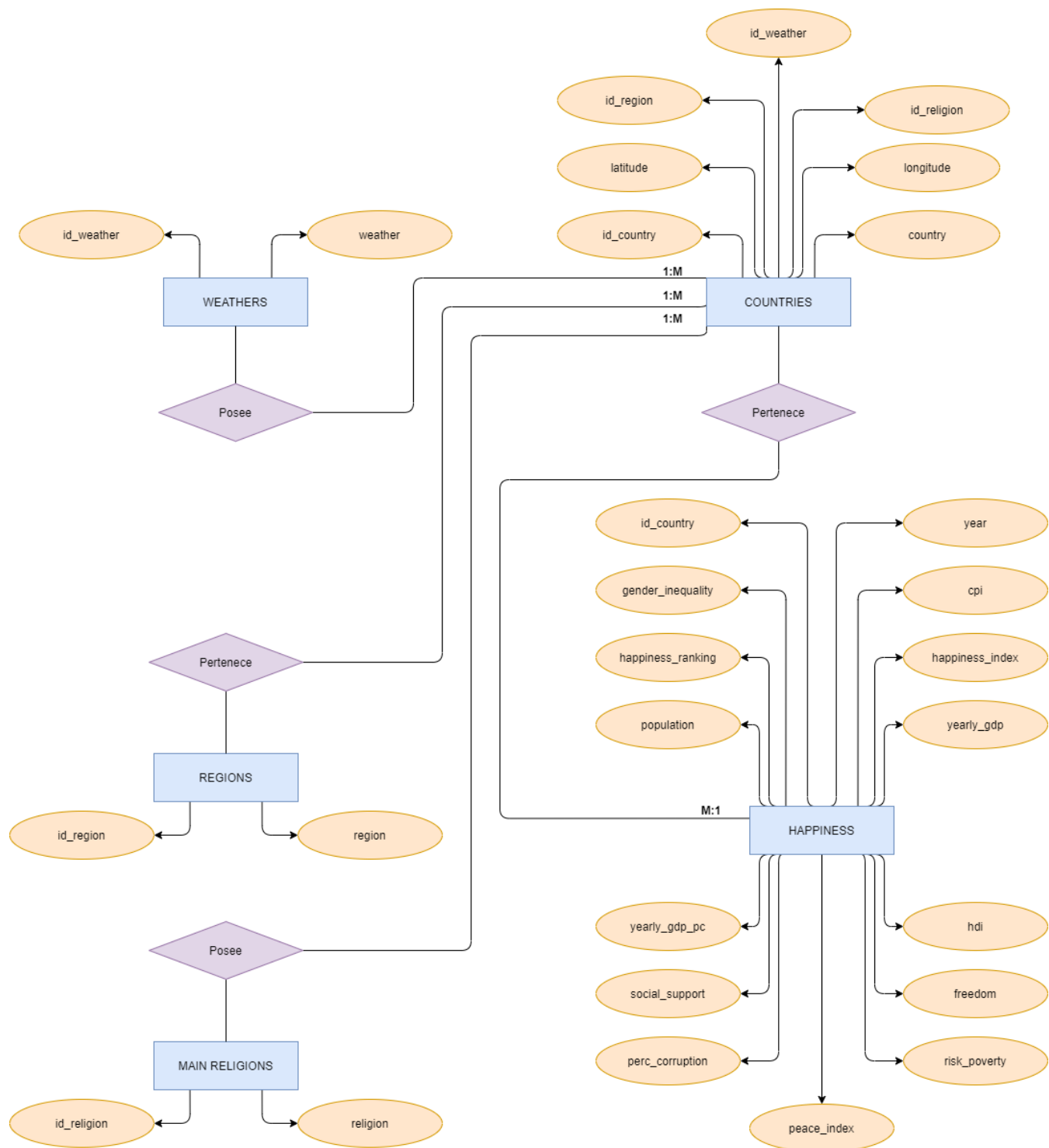
Longitud (longitude): Longitud de la capital del país.

Región (region): Región geográfica a la que pertenece el país.

Religión (religion): Religión con mayor porcentaje de creyentes de un país.

Clima (weather): Clima aproximado de un país según temperaturas medias del mismo.

DIAGRAMAS ENTIDAD-RELACION



INTEGRANTES:

Enzo Matías González
Egisto Barazarte González

pág. 5

DEFINICION DE TIPO DE CAMPO

HAPPINESS

| TIPO DE CLAVE | CAMPO | TIPO DE CAMPO |
|---------------|---------------------------|---------------|
| PK | id_happiness | int |
| FK | id_country | int |
| - | year | int |
| - | happiness_ranking | int |
| - | happiness_index | decimal(6,3) |
| - | population | int |
| - | yearly_gdp | int |
| - | yearly_gdp_pc | decimal(32,8) |
| - | hdi | decimal(6,3) |
| - | social_support | decimal(6,3) |
| - | freedom | decimal(6,3) |
| - | perceptions_of_corruption | decimal(6,3) |
| - | poverty | decimal(6,3) |
| - | gender_inequality | decimal(6,3) |
| - | cpi | decimal(10,3) |
| - | peace_index | decimal(6,3) |

COUNTRIES

| TIPO DE CLAVE | CAMPO | TIPO DE CAMPO |
|---------------|-------------|---------------|
| PK | id_country | int |
| FK | id_region | int |
| FK | id_weather | int |
| FK | id_religion | int |
| - | country | nvarchar(255) |
| - | latitude | nvarchar(255) |
| - | longitude | nvarchar(255) |

INTEGRANTES:

Enzo Matías González
Egisto Barazarte González

DEFINICION DE TIPO DE CAMPO

REGION

| TIPO DE CLAVE | CAMPO | TIPO DE CAMPO |
|---------------|-----------|---------------|
| PK | id_region | int |
| - | region | nvarchar(255) |

RELIGION

| TIPO DE CLAVE | CAMPO | TIPO DE CAMPO |
|---------------|-------------|---------------|
| PK | id_religion | int |
| - | religion | nvarchar(255) |

WEATHERS

| TIPO DE CLAVE | CAMPO | TIPO DE CAMPO |
|---------------|------------|---------------|
| PK | id_weather | int |
| - | weather | nvarchar(255) |

TRANSFORMACIONES DE DATOS

- Se quitaron filas en blanco de la tabla de "Happiness"

```
##Filas en blanco eliminadas" = Table.SelectRows("#Tipo cambiado", each not  
List.IsEmpty(List.RemoveMatchingItems(Record.FieldValues(_), {"", null})))
```

- Se cambió el nombre de la columna "Countries" a "Name" en la tabla llamada "Countries"

```
##Columnas con nombre cambiado" = Table.RenameColumns("#Valor reemplazado",{"Countries", "Name"})
```

- Se cambió la "," que separaba en los nombres en la columna "Name" en la tabla llamada "Countries"

```
##Valor reemplazado" = Table.ReplaceValue("#Columnas con nombre cambiado", ",", " ", Replacer.ReplaceText, {"Name"})
```

- Se cambió el tipo de dato de las columnas "Yearly GDP (M U\$S)" y "Yearly GDP Per Capita" de Número decimal a Número decimal fijo en la tabla "Happiness"

```
##Tipo cambiado1" = Table.TransformColumnTypes("#Filas en blanco eliminadas",{"Yearly GDP (M U$S)",  
Currency.Type}, {"Yearly GDP Per Capita", Currency.Type}))
```

- Se creó una nueva tabla que contiene solo la columna "Year" de la tabla de "Happiness" y se la nombró "Fecha"

```
Origen = Excel.Workbook(File.Contents("C:\Users\dgenz\Desktop\Data Analyst\Proyecto  
Final\Dataset\Felicidad(Dataset_Separado).xlsx"), null, true),
```

```
Happiness_Sheet = Origen[Item="Happiness",Kind="Sheet"]][Data],
```

```
##Encabezados promovidos" = Table.PromoteHeaders(Happiness_Sheet, [PromoteAllScalars=true]),
```

```
##Tipo cambiado" = Table.TransformColumnTypes("#Encabezados promovidos",{"ID_Happyness", Int64.Type,  
{"Year", Int64.Type}, {"ID_Country", Int64.Type}, {"Happiness Ranking", Int64.Type}, {"Happiness Index", type  
number}, {"Population", Int64.Type}, {"Yearly GDP (M U$S)", type number}, {"Yearly GDP Per Capita", type number},  
{"HDI", type number}, {"Social support", type number}, {"Freedom to make life choices", type number},  
{"Perceptions of corruption", type number}, {"% Risk of poverty", type number}, {"Gender inequality", type number},  
{"CPI", type number}, {"Peace Index", type number})),
```

```
##Filas en blanco eliminadas" = Table.SelectRows("#Tipo cambiado", each not  
List.IsEmpty(List.RemoveMatchingItems(Record.FieldValues(_), {"", null}))),
```

```
##Otras columnas quitadas" = Table.SelectColumns("#Filas en blanco eliminadas",{"Year"})
```

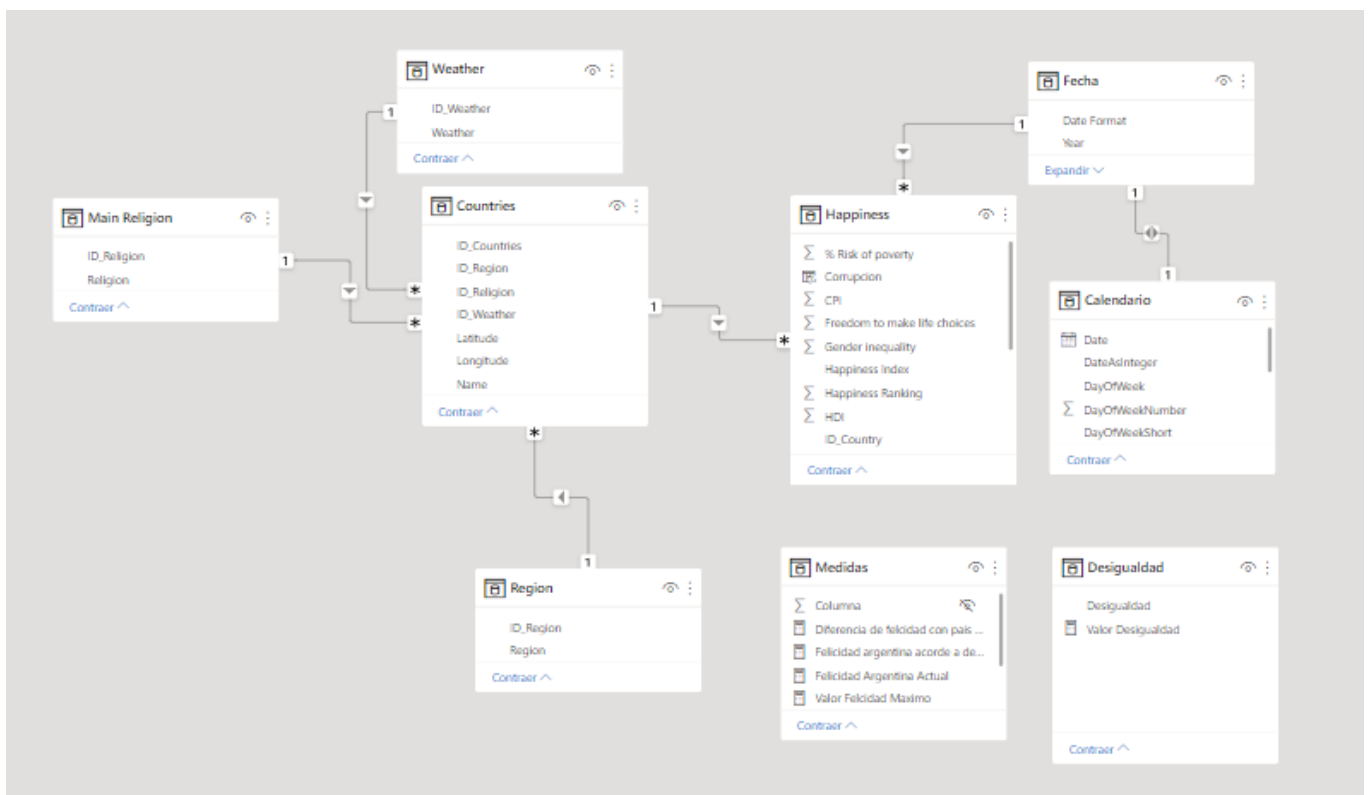
- Se eliminaron los duplicados de en la columna "Year" en la tabla "Fecha"

```
##Duplicados quitados" = Table.Distinct("#Otras columnas quitadas")
```

- Se agregó una columna en la tabla llamada "Fecha" llamada "Date Format" la cual da formato de fecha a los datos en la columna "Year"

```
##Fecha insertada" = Table.AddColumn("#Duplicados quitados", "Date Format", each  
Date.From(Text.From([Year], "es-AR")), type date)
```

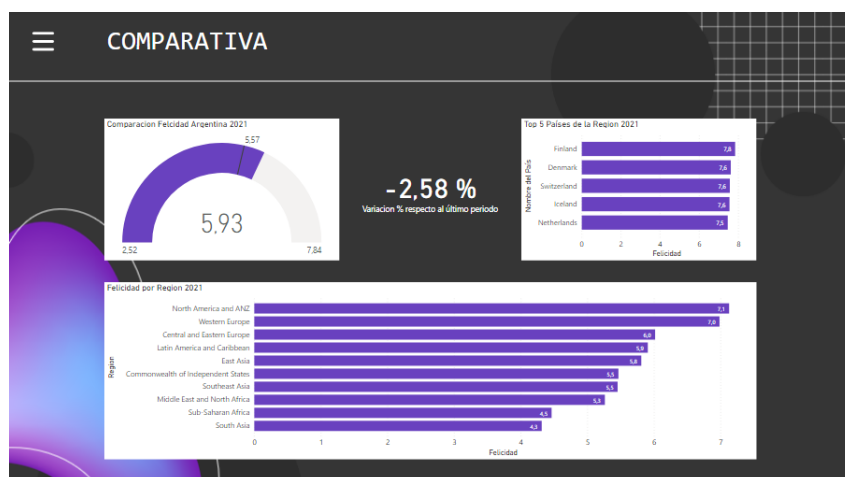

DIAGRAMA RELACIONAL DEL TABLERO



ANALISIS FUNCIONAL DEL TABLERO



PORTADA: Introduce al usuario al proyecto desde donde este mismo podrá visualizar un mapa interactivo con los países analizados y al hacer click en ellos para poder ver la felicidad actual(2021) de los mismos. Además, a la derecha tenemos diferentes tarjetas de donde se puede encontrar información del alcance del análisis realizado. A su vez, al lado izquierdo se encuentra el panel de navegación.



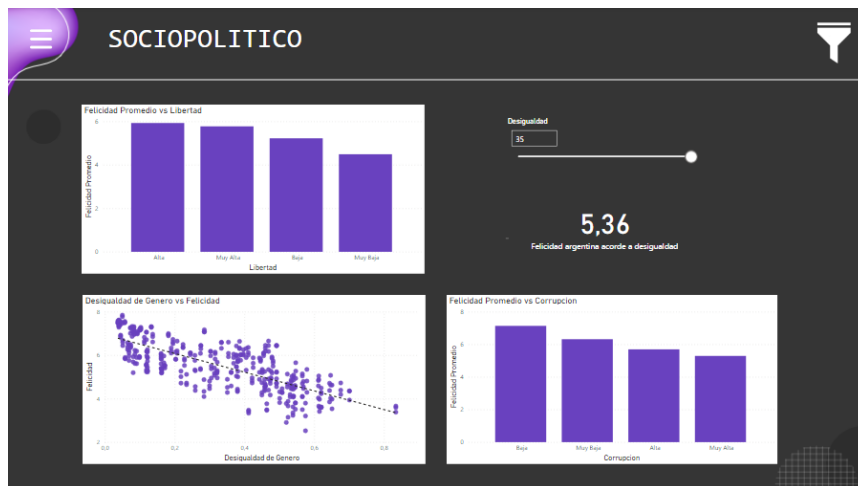
COMPARATIVA: Tal como dice el nombre en esta solapa podemos realizar una comparación entre la Argentina y el resto del mundo o de la región que seleccione el usuario en el gráfico de “Felicidad por Región 2021”.

“Comparación Felicidad Argentina 2021” muestra la felicidad actual argentina en el centro y toma como valor mínimo el peor país de la región considerada y como máximo a la mejor mientras que el valor objetivo es el promedio de la región.

ANALISIS FUNCIONAL DEL TABLERO

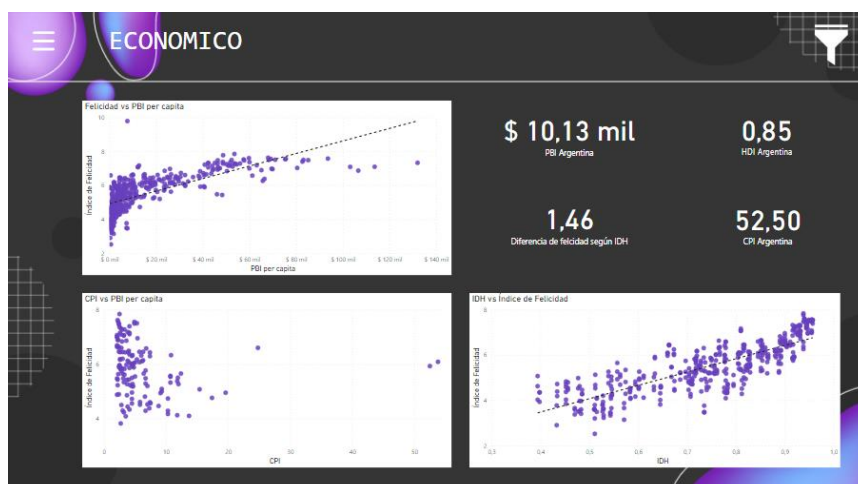
“Top 5 países de la Región” muestra los 5 países con mayor felicidad del mundo o de la región seleccionada.

“Felicidad por Región 2021” muestra la felicidad promedio de las diferentes regiones del mundo analizadas, desde aquí se puede seleccionar una región para usarla como filtro para los otros dos gráficos.



SOCIOPOLITICO: En esta solapa se analizan aspectos de índole político y como afectan a los índices de felicidad. En esta solapa se observan dos gráficos de barras que comparan el nivel promedio de felicidad según el grado de libertad y según el nivel de corrupción que perciben los ciudadanos de un país.

Por otro lado, se busca ver la influencia de la Desigualdad de Género que posee el país sobre la felicidad. En la esquina superior derecha de la solapa se agregaron dos segmentadores de datos para un análisis más detallado por región y año.



INTEGRANTES:

Enzo Matías González
Egisto Barazarte González

ANALISIS FUNCIONAL DEL TABLERO

ECONOMICO: En esta solapa se analizan los aspectos de índole económico y como afectan a los índices de felicidad.

Se observan tres gráficos que buscan relacionar el PBI per cápita, el Índice de Precios al Consumidor y el Índice de Desarrollo Humano con el nivel de felicidad. En la esquina superior derecha de la solapa se agregaron dos segmentadores de datos para un análisis más detallado por región y año, además de cuatro etiquetas con los índices que posee la argentina en estas áreas



CLIMATICO Y CULTURAL: Como dice el nombre de la solapa se busca analizar conjuntamente el promedio de felicidad dividido por religión y por clima, además se agregó una grafico que muestra la pobreza en función del clima como una posible causa de su efecto en la felicidad.

MEDIDAS CALCULADAS

- Valor Felicidad Promedio
 - Compuesta por:
Happiness[Happiness Index]
 - Finalidad: Se usa para obtener el valor promedio del índice de felicidad de una región para usarlo en un gráfico de medidor tipo velocímetro.
 - Fórmula: AVERAGE(Happiness[Happiness Index])
- Valor Felicidad Máximo
 - Compuesta por:
Happiness[Happiness Index]
 - Finalidad: Se usa para obtener el mayor índice de felicidad para usarlo en un gráfico de medidor tipo velocímetro.
 - Fórmula: MAX(Happiness[Happiness Index])
- Valor Felicidad Mínimo
 - Compuesta por:
Happiness[Happiness Index]
 - Finalidad: Se usa para obtener el menor índice de felicidad para usarlo en un gráfico de medidor tipo velocímetro.
 - Fórmula: MIN(Happiness[Happiness Index])
- Felicidad Argentina Actual
 - Compuesta por:
Happiness[Happiness Index]
Countries[Name]
Calendario[Year]
Region[Region]
 - Finalidad: Se usa para obtener el ultimo índice de felicidad registrado de la Argentina para usarlo en un gráfico de medidor tipo velocímetro.
 - Fórmula: CALCULATE(SUM(Happiness[Happiness Index]), Countries[Name] = "Argentina", Calendario[Year]=2021, Region[Region] = "Latin America and Caribbean")
- Felicidad argentina acorde a desigualdad
 - Compuesta por:
Desigualdad[Valor Desigualdad]
Happiness[Gender inequality]
Happiness[Happiness Index]
 - Finalidad: Se utiliza junto a un parámetro para dar una estimación de cuanto sería el índice de felicidad si se modificara dicho parámetro.
 - Fórmula: CALCULATE(Desigualdad[Valor Desigualdad]*(-0.03)+ CALCULATE(AVERAGE(Happiness[Happiness Index]), Happiness[Gender inequality]<0.09))

MEDIDAS CALCULADAS

- Diferencia de felicidad con país de mayor IDH
 - Compuesta por:
Happiness[ID_Country]
Happiness[HDI]
Happiness[Happiness Index]
Calendario[Year]
[Felicidad Argentina Actual]
 - Finalidad: Se busca mostrar la brecha entre la argentina y el país con mayor IDH para remarcar la importancia del mismo
 - Fórmula: `VAR mejor_pais = LOOKUPVALUE(Happiness[ID_Country], Happiness[HDI], MAX(Happiness[HDI]))`
`VAR felicidad_m_idh = CALCULATE(SUM(Happiness[Happiness Index]), Happiness[ID_Country] = mejor_pais, Calendario[Year]=2021)`
`RETURN CALCULATE(felicidad_m_idh-[Felicidad Argentina Actual])`
- Variación % respecto a la medida anterior
 - Compuesta por:
Happiness[Happiness Index]
Calendario[Date]
Calendario[Year]
Happiness[ID_Country]
[Felicidad Argentina Actual]
 - Finalidad: Se usa para obtener el valor promedio del índice de felicidad de una región para usarlo en un gráfico de medidor tipo velocímetro.
 - Fórmula: `VAR anio = YEAR(LASTDATE(DATEADD(Calendario[Date],-2,year)))`
`VAR media = CALCULATE(SUM(Happiness[Happiness Index]), Calendario[Year]=anio, Happiness[ID_Country]=4)`
`RETURN CALCULATE((([Felicidad Argentina Actual]-media)/media)`

HERRAMIENTAS UTILIZADAS Y FUTURAS LINEAS

HERRAMIENTAS UTILIZADAS

En este proyecto se usó Excel para ejecutar y realizar una primera “limpieza” del archivo .csv en el que se encontraba el dataset utilizado

Mientras que por otro lado utilizamos Power BI para crear el dashboard donde se realizó el análisis del dataset .

FUTURAS LINEAS

Consideramos este proyecto una oportunidad muy grande para indagar mas en como se relacionan los indicadores estudiados con el nivel de felicidad de un país, particularmente como estos influyen también a los otros indicadores, ya que estos generan un efecto cascada que desencadena en una gran fluctuación en la felicidad.

Por otro lado, consideramos que se podría indagar en la posibilidad de hacer proyecciones aproximadas de como afectaría el variar estos indicadores y a su vez buscara predecir la dirección en la que evolucionan las necesidades de las personas para poder tomar decisiones a largo plazo. Para esto sería importante el poseer más información de años anteriores al igual que poder agregar más indicadores de índole social ya que los mismos demostraron ser los mas influyentes en la felicidad de las personas.

Otra oportunidad seria directamente acotar el área de estudio a únicamente de índole sociopolítico ya que además de ser el más influyente también es el que es más posible de modificar al contrario del clima, la cultura, etc.