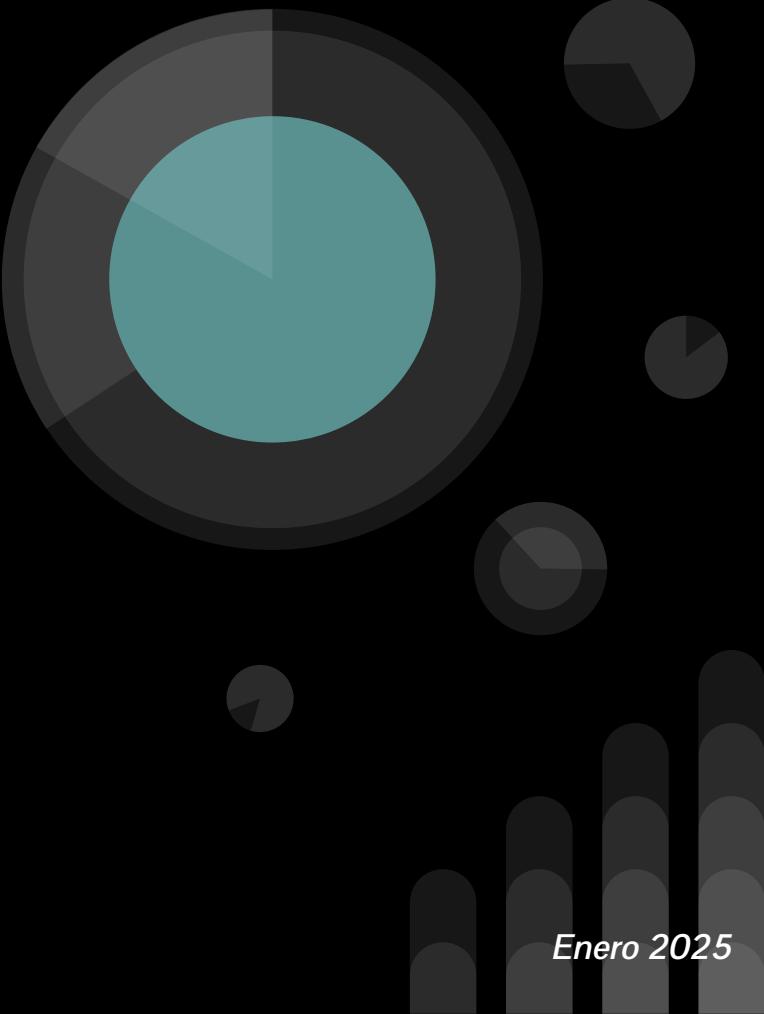


Project The Bridge

EDA _ N0m4d1sm0 digital



Miembros Team 4 : Daniela Aguirre, Alejandro Balaguer y Juan F. Cía

Enero 2025

Introducción: contexto de negocio



Qué buscamos con el análisis

El nomadismo digital ha crecido de forma acelerada, pero no todos los destinos son igual de atractivos para este perfil. Elegir un destino depende de la evaluación de varios factores entre lo que tenemos:

- **Coste de vida:** vivienda, alimentación y transporte.
- **Infraestructura digital:** velocidad y coste de internet.
- **Calidad de vida:** seguridad, bienestar y servicios.
- **Capacidad económica local:** relación entre salarios y el coste mensual de productos y servicios.
- **Accesibilidad burocrática a trabajar en esos destinos:** facilidad de visa y requisitos legales.



Impacto en negocio

Este tipo de análisis puede tener interés no solo para los propios nómadas digitales, también para gobiernos, instituciones o empresas del tejido productivo para tomar decisiones.

- **Gobiernos e instituciones:** atraer profesionales cualificados e inversiones.
- **Empresas remote-first:** oportunidad competitiva para obtener talento profesional.
- **Empresas del mundo inmobiliario:** los nómadas buscan viviendas en países de destino.
- **Empresas de tramitaciones burocráticas:** tramitación de documentación legal.
- **Plataformas de relocación:** facilitación del movimiento entre países y empresas.
- **Empresas de servicios específicos para nomadismo.**

Objetivo del EDA

¿Qué características económicas y de coste de vida
diferencian a los países y ciudades mejor posicionados
para el nomadismo digital frente a los menos atractivos?



Hipótesis a validar en el análisis

El análisis se estructura en torno a hipótesis primarias y secundarias:

Hipótesis primarias

- HIPÓTESIS 1

Coste de vida medio: ¿los mejores destinos siempre son los que menos coste mensual exigen?

- HIPÓTESIS 2

Salarios vs coste: ¿la relación entre los salarios y el coste es clave para ser un buen destino?

- HIPÓTESIS 3

Conectividad: ¿cuanto más velocidad de conexión a menos precio, mejor destino nómada?



Hipótesis a validar en el análisis

El análisis se estructura en torno a hipótesis primarias y secundarias:

Hipótesis secundarias

- HIPÓTESIS 4

Vivienda: ¿el alquiler de vivienda es una razón determinante por encima de alimentación?

- HIPÓTESIS 5

Más por menos dinero: ¿hay países con mejores condiciones por menos coste mensual?

- HIPÓTESIS 6

Bienestar social: ¿los países con buen índice de felicidad atraen más nómadas digitales?



Hipótesis a validar en el análisis

El análisis se estructura en torno a hipótesis primarias y secundarias:

Hipótesis secundarias

- HIPÓTESIS 7

Outliers: ¿países fuera de rango por coste y puntuación que son buenos destinos?

- HIPÓTESIS 8

Seguridad: ¿la seguridad es un elemento decisivo en la elección, incluso con costes elevados?

- HIPÓTESIS 9

Facilidad de visado: ¿la accesibilidad legal a los países afecta al ranking de destinos?

Descripción de los datasets

Dataset de variables originales y agregadas de Cost of Living

Número de filas: 4.742 ciudades

Número de variables: 65 en total (58 originales + 6 variables agregadas)

Fuente: Numbeo / Kaggle

<class 'pandas.core.frame.DataFrame'>			
RangeIndex: 4956 entries, 0 to 4955			
Data columns (total 58 columns):			
#	Column	Non-Null Count	Dtype
---	---	-----	----
0	city_name	4956 non-null	object
1	country_name	4956 non-null	object
2	meal_inexpensive_restaurant	4528 non-null	float64
3	meal_midrange_restaurant_2p	4505 non-null	float64
4	mcmeal_fastfood	4622 non-null	float64
5	beer_domestic_restaurant_0_5l	4460 non-null	float64
6	beer_imported_restaurant_0_33l	4516 non-null	float64
7	cappuccino_restaurant	4612 non-null	float64
8	soda_restaurant_0_33l	4511 non-null	float64
9	water_restaurant_0_33l	4640 non-null	float64
10	milk_1l	4578 non-null	float64
11	bread_white_500g	4543 non-null	float64
12	rice_white_1kg	4613 non-null	float64
13	eggs_12	4449 non-null	float64
14	cheese_local_1kg	4478 non-null	float64
15	chicken_fillet_1kg	4398 non-null	float64
16	beef_1kg	4390 non-null	float64
17	apples_1kg	4584 non-null	float64
18	bananas_1kg	4575 non-null	float64
19	oranges_1kg	4515 non-null	float64
...			
56	mortgage_interest_rate_20y	3950 non-null	float64
57	data_quality_flag	4956 non-null	int64
dtypes: float64(55), int64(1), object(2)			
memory usage: 2.2+ MB			



Dataset de variables originales y agregadas de Cost of Living

Número de filas: 4.742 ciudades

Número de variables: 65 en total (58 originales + 7 variables agregadas)

Fuente: Numbeo / Kaggle

Variables agregadas

- Promedio del alquiler de un dormitorio (centro + afueras): *nomad_housing_cost*.
- Promedio de productos básicos de supermercado: *basic_basket_index*.
- Promedio de gasto diario en comer fuera (cappuccino + comida de restaurante más económico): *daily_meal_cost*.
- Coste de vida mensual (costes en vivienda + comida + internet + utilities + transporte): *monthly_nomad_cost*.
- Poder adquisitivo (salario / coste mensual nómada): *local_purchasing_power*.
- Porcentaje de salario destinado a alquiler de vivienda: *housing_salary_ratio*.



Dataset de Nomad Index de Circleloop

Número de filas: 85 países

Número de variables: 10 variables

Fuente: Circleloop

Variables propias de este dataset

- Conectividad a internet general y móvil: *broadband_speed_mbps* y *mobile_speed_mbps*.
- Costes en conectividad y alquiler: *broadband_cost* y *monthly_rent*.
- Índice de bienestar: *happiness_index*.
- Variable clave con la puntuación sobre mejores destino para nómadas digitales: *digital_nomad_score*.



Dataset de Nomad Index de Movingto

Número de filas: 40 registros

Número de variables: 10 variables

Fuente: Movingto

Variables propias de este dataset

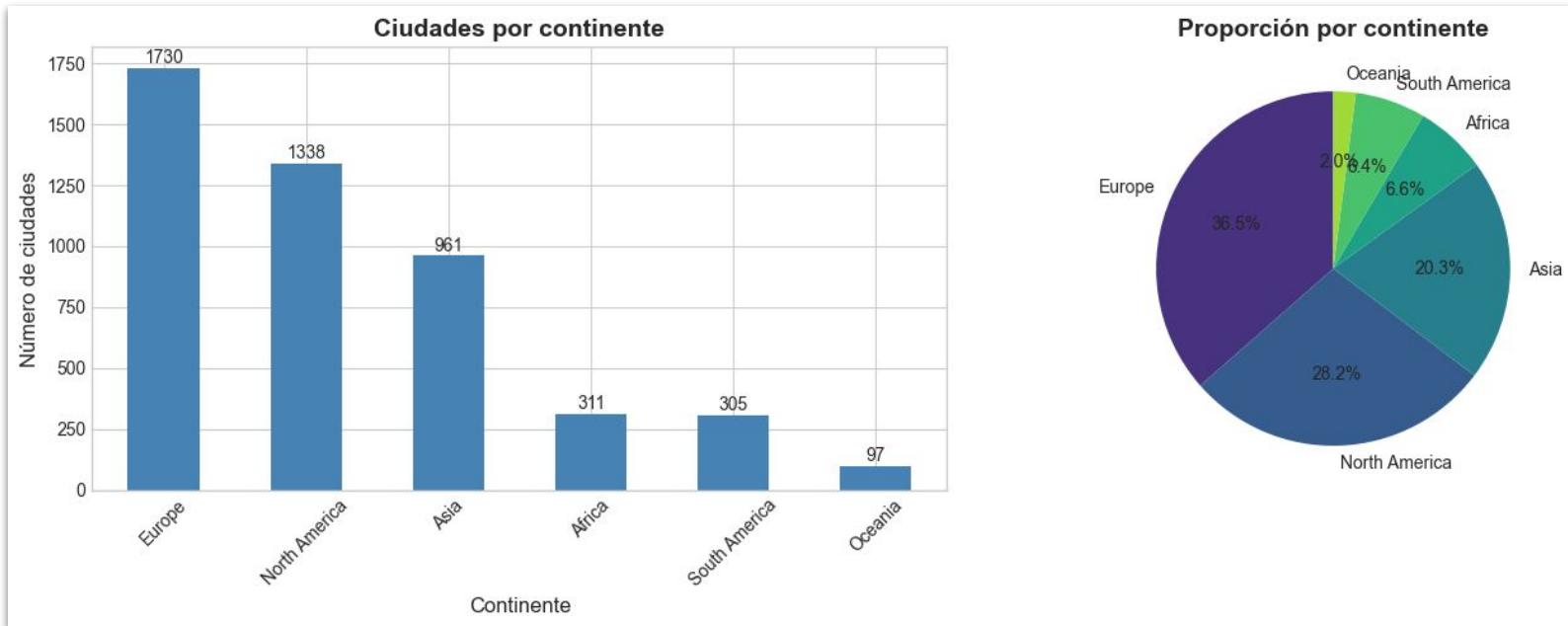
- Score de 0 a 10 en seguridad: `safety`.
- Facilidad de obtención de visa para nómadas digitales de 0 a 100: `visa_ease`.
- Información fiscal sobre impuestos: `taxes` y `tax_free_period`.

Análisis univariante

[Dataset de Cost of Living]

Variables categóricas [Cost of Living]

Volumen de ciudades por continente





Top 10 de países con más ciudades

Top 10 de países con más ciudades:

country_name	ciudades
United States	1027
Italy	170
India	165
Brazil	157
United Kingdom	155
Germany	144
Russia	129
Canada	117
France	99
China	99

Variables numéricas originales [Cost of Living]



Variables originales [Cost of Living]

Valores estadísticos descriptivos de las variables originales de Cost of Living:

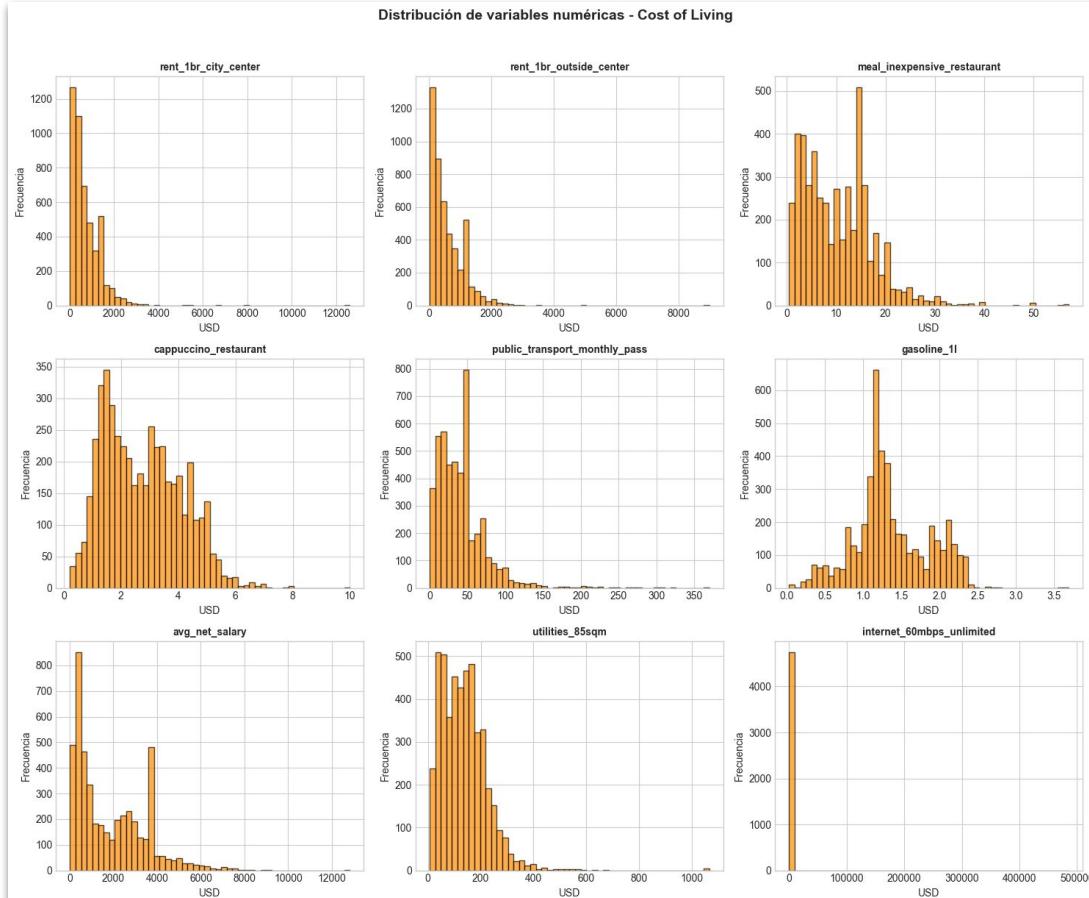
	rent_1br_city_center	rent_1br_outside_center	meal_inexpensive_restaurant	cappuccino_restaurant	public_transport_monthly_pass	gasoline_1l
count	4742.00	4742.00	4742.00	4742.00	4742.00	4742.00
mean	707.64	565.72	10.39	2.74	40.79	1.36
std	620.69	505.84	7.02	1.35	31.45	0.49
min	21.70	12.20	0.45	0.22	0.00	0.03
25%	258.98	177.94	4.48	1.58	17.79	1.10
50%	526.87	410.96	10.00	2.62	36.88	1.26
75%	1048.31	850.00	15.00	3.75	50.00	1.71
max	12608.83	8989.37	57.14	10.00	369.86	3.70



Variables numéricas originales, no agregadas de Cost of Living

1. Analizamos variables originales que son relevantes para un nómada: vivienda, alimentación, transporte y salarios.

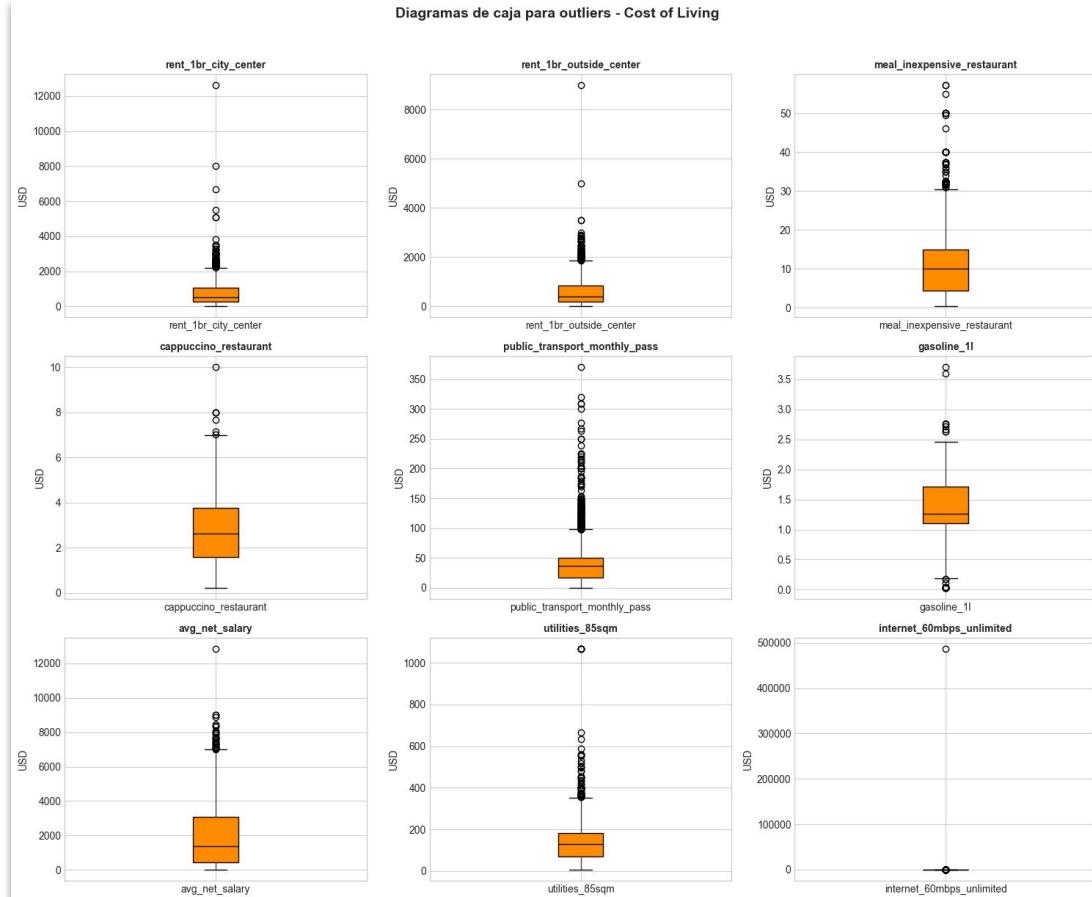
2. Mostramos distribuciones con histogramas y valores atípicos con diagramas de caja.



Variables numéricas originales, no agregadas de Cost of Living

1. Analizamos variables originales que son relevantes para un nómada: vivienda, alimentación, transporte y salarios.

2. Mostramos distribuciones con histogramas y valores atípicos con diagramas de caja.



Variables numéricas agregadas [Cost of Living]



Variables agregadas [Cost of Living]

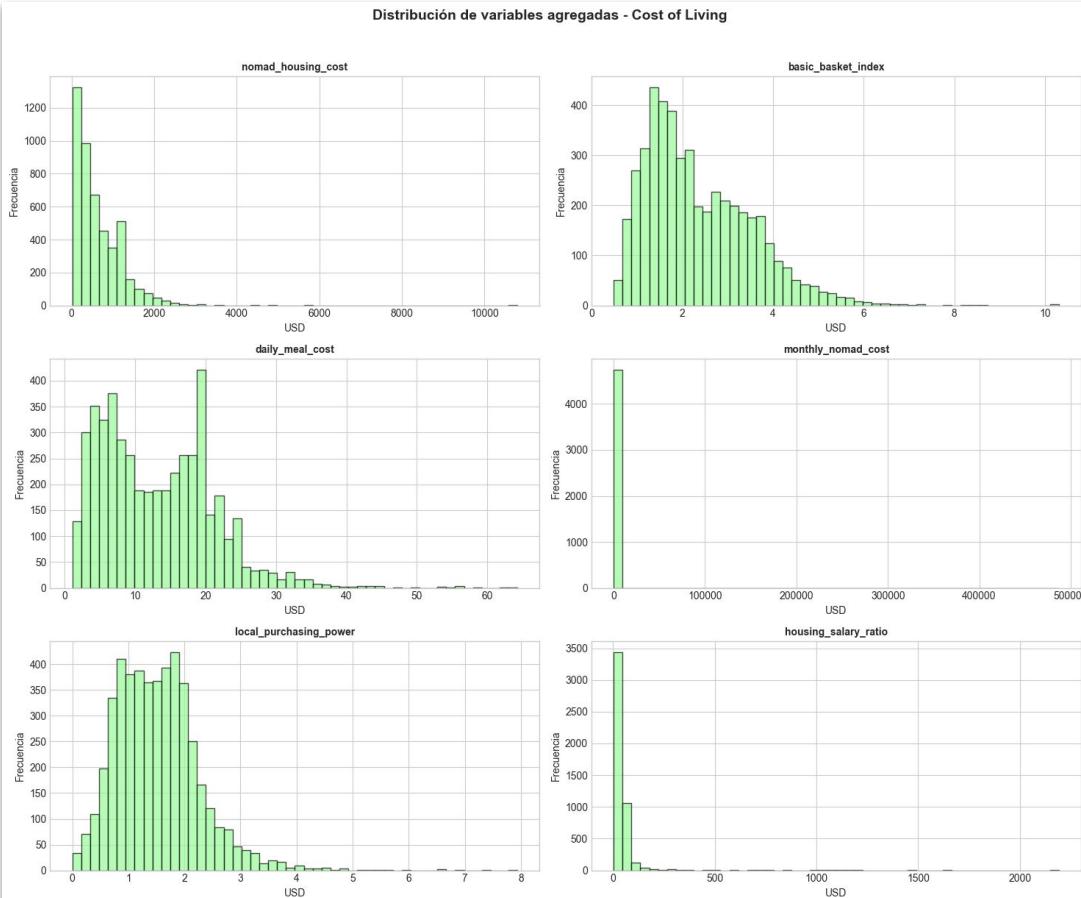
Valores estadísticos descriptivos de las variables agregadas de Cost of Living:

	nomad_housing_cost	basic_basket_index	daily_meal_cost	monthly_nomad_cost	local_purchasing_power	housing_salary_ratio
count	4742.00	4742.00	4742.00	4742.00	4742.00	4742.00
mean	636.68	2.35	13.14	1223.50	1.52	47.39
std	553.18	1.17	7.96	7094.61	0.75	75.85
min	18.98	0.49	1.10	108.08	0.00	2.94
25%	219.81	1.44	6.37	491.88	0.97	26.95
50%	471.72	2.08	12.26	967.96	1.46	34.68
75%	945.63	3.14	18.92	1642.79	1.94	49.12
max	10799.10	10.30	64.28	486948.09	7.94	2191.42



Variables numéricas agregadas de Cost of Living

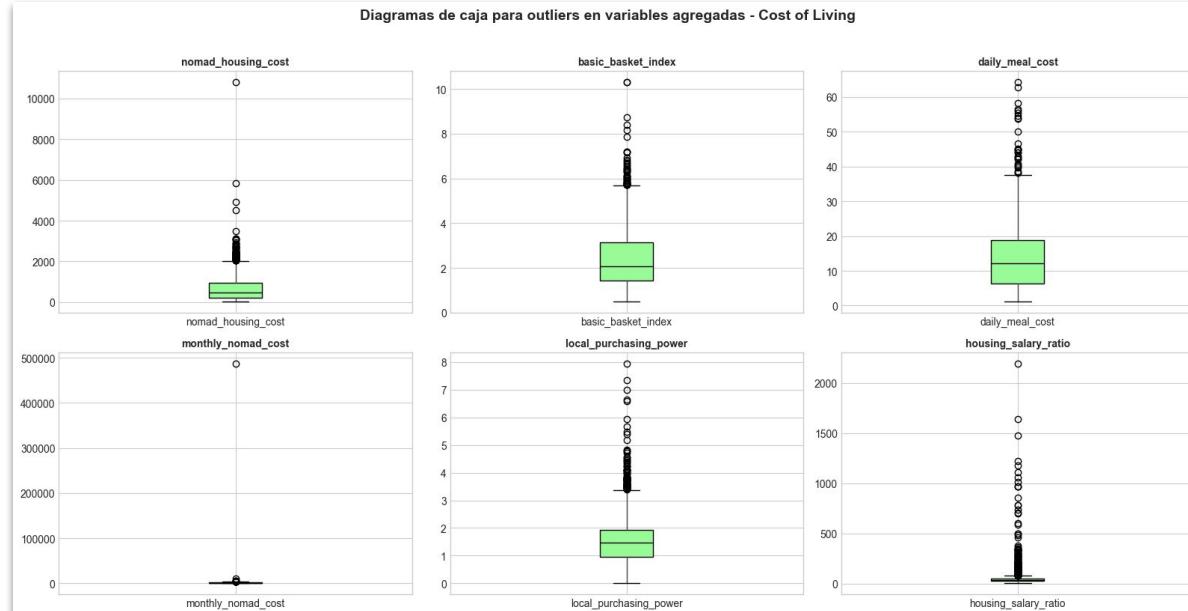
- Mostramos distribuciones con histogramas y valores atípicos con diagramas de caja.
- Análisis porcentual de variables agregadas de Cost of Living:
 1. Análisis por percentiles del coste medio mensual en todas las ciudades.
 2. Capacidad de compra por encima de salario medio.
 3. Volumen de inversión del salario medio en vivienda.





Variables numéricas agregadas de Cost of Living

- Mostramos distribuciones con histogramas y valores atípicos con diagramas de caja.
- Análisis porcentual de variables agregadas de Cost of Living:
 1. Análisis por percentiles del coste medio mensual en todas las ciudades.
 2. Capacidad de compra por encima de salario medio.
 3. Volumen de inversión del salario medio en vivienda.





Análisis de ciudades baratas, moderadas, caras y muy caras en función de los percentiles del gasto mensual

Coste mensual de la vida en todas las ciudades del dataset Cost of Living

El análisis establece los cuartiles (25% y 75%), la mediana y el percentil 90

Coste mensual barato (percentil 25): 492 dólares

Coste mensual moderado (percentil 50): 968 dólares

Coste mensual caro (percentil 75): 1643 dólares

Coste mensual muy caro (percentil 90): 2113 dólares



Capacidad de compra mensual: ¿qué % de ciudades tienen salarios promedio por encima del coste mensual?

Análisis de capacidad de compra (salario promedio/coste mensual)

Ciudades donde salario > coste mensual: 3494 (73.7%)

Ciudades donde salario < coste mensual: 1248 (26.3%)

Los ciudadanos del 73,7% de las ciudades ingresan un salario promedio superior al gasto mensual

Los ciudadanos del 26,3% de las ciudades ingresan un salario promedio inferior al gasto mensual



Capacidad de compra mensual: ¿qué % de ciudades tienen salarios promedio por encima del coste mensual?

Análisis del gasto porcentual del salario promedio en vivienda

Situación óptima (gasto < 30%): 1591 ciudades (33.6%)

Situación mejorables (gasto entre 30-50%): 2014 ciudades (42.5%)

Situación arriesgada (gastos > 50%): 1137 ciudades (24.0%)

[Dataset de Nomad Index Circleloop]



Variables numéricas [Circleloop]

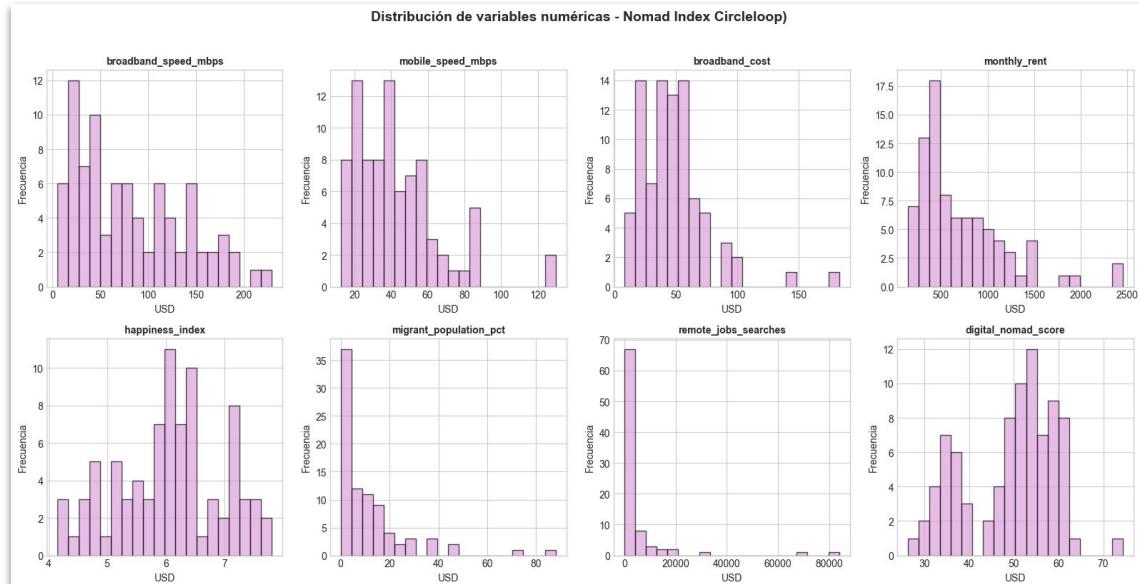
Valores estadísticos descriptivos de las variables de Nomad Index Cicleloop:

	rank	broadband_speed_mbps	mobile_speed_mbps	broadband_cost	monthly_rent	happiness_index	migrant_population_pct
count	85.00	85.00	85.00	85.00	85.00	85.00	85.00
mean	43.00	81.16	42.68	47.99	712.93	6.06	11.51
std	24.68	56.41	23.16	27.75	470.69	0.89	15.39
min	1.00	4.97	12.68	8.34	145.08	4.15	0.10
25%	22.00	31.89	24.63	31.46	382.59	5.51	1.70
50%	43.00	68.90	38.88	45.18	558.09	6.12	5.30
75%	64.00	120.13	53.89	59.72	943.02	6.66	15.40
max	85.00	229.42	129.61	183.17	2452.32	7.81	87.90

Variables numéricas de Index Circleloop

Mostramos distribuciones con histogramas y valores atípicos con diagramas de caja.

1. Son relevantes las distribuciones de la velocidad de conexión a internet y el índice de bienestar.
2. Esas dos variables son posiblemente las más relevantes en el score positivos en los índices de destino para nómadas digitales.
3. Y vemos que tienen distribuciones de los datos muy diferentes.

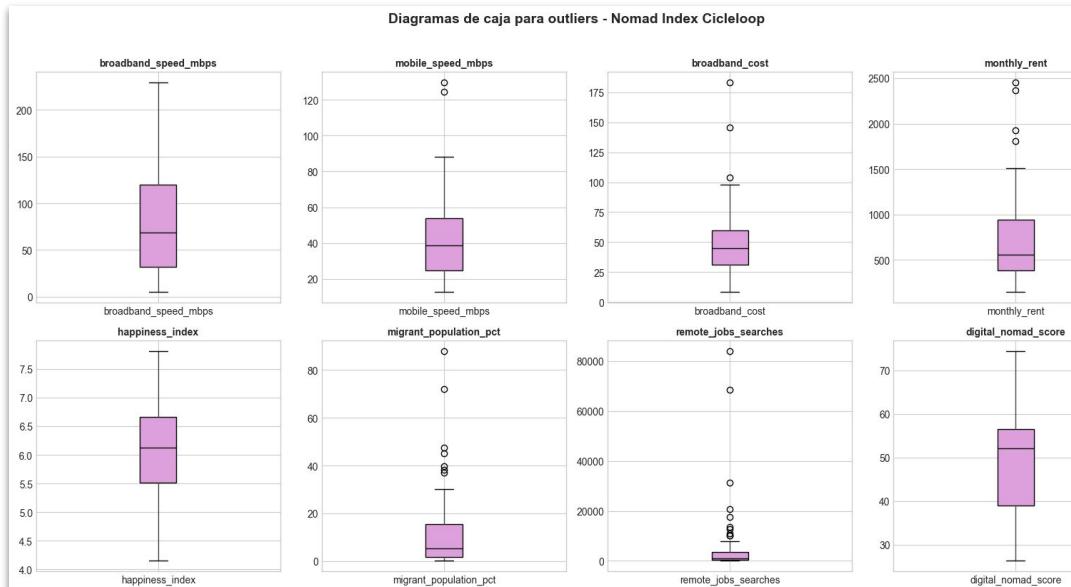




Variables numéricas de Index Circleloop

Mostramos distribuciones con histogramas y valores atípicos con diagramas de caja.

1. En los diagramas de caja se puede ver que las dos variables más relevantes no tienen outliers.





Top 10 países con mejor puntuación en destinos para nómadas digitales

Top 10 países con mejor puntuación en destinos para nómadas digitales:

country	digital_nomad_score
Canada	74.35
UK	63.43
Romania	62.28
Sweden	61.54
Denmark	61.49
France	60.80
Netherlands	60.27
Australia	60.16
Switzerland	60.15
Germany	60.00



Top 10 países con peor puntuación en destinos para nómadas digitales

Top 10 países con peor puntuación en destinos para nómadas digitales:

country	digital_nomad_score
Nigeria	26.32
Myanmar	29.16
Ghana	30.36
Kenya	31.27
Sri Lanka	32.08
Cambodia	32.82
Algeria	33.38
Egypt	33.84
Tunisia	34.15
Honduras	34.48

[Dataset de Nomad Index Movingto]

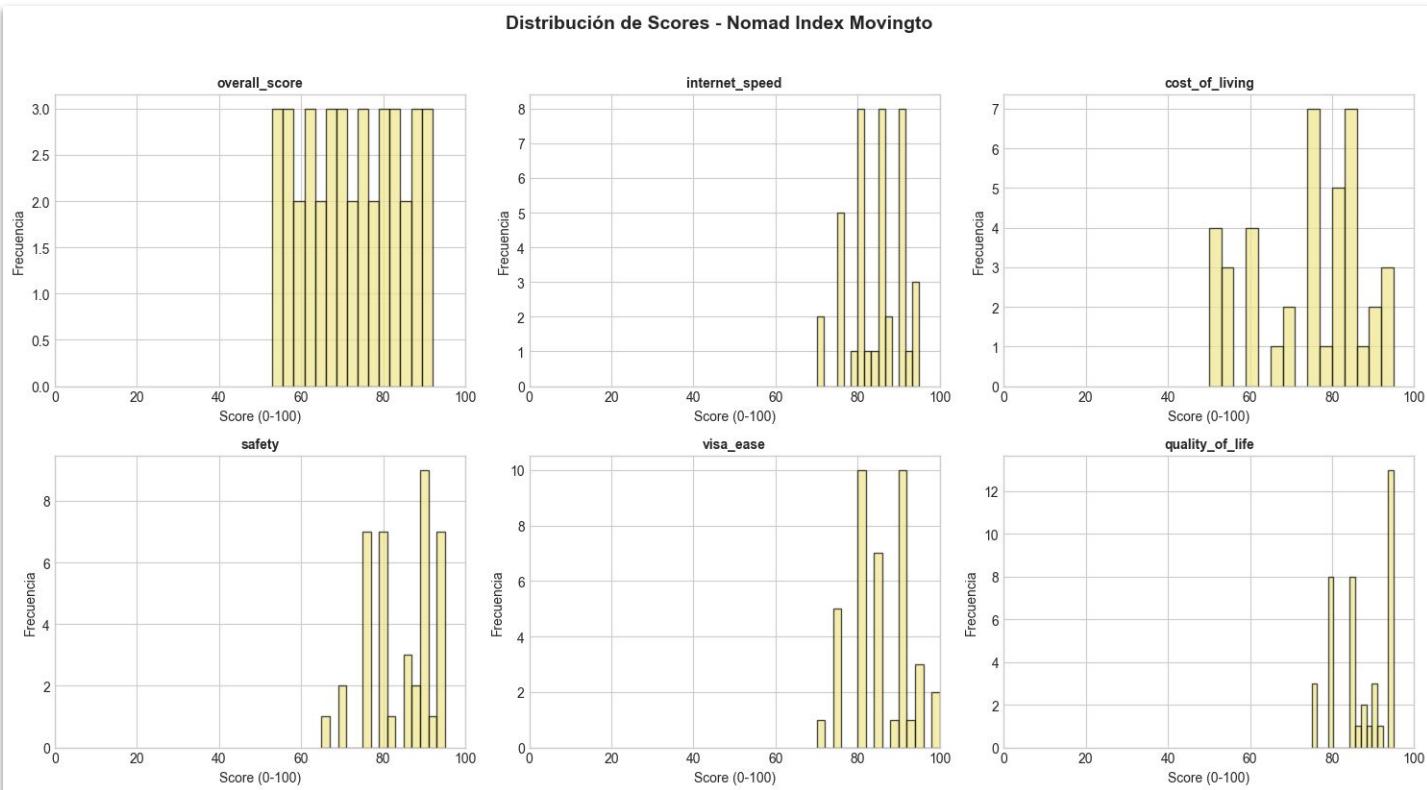
Variables numéricas [Movingto]

Valores estadísticos descriptivos de las variables de Nomad Index Movingto:

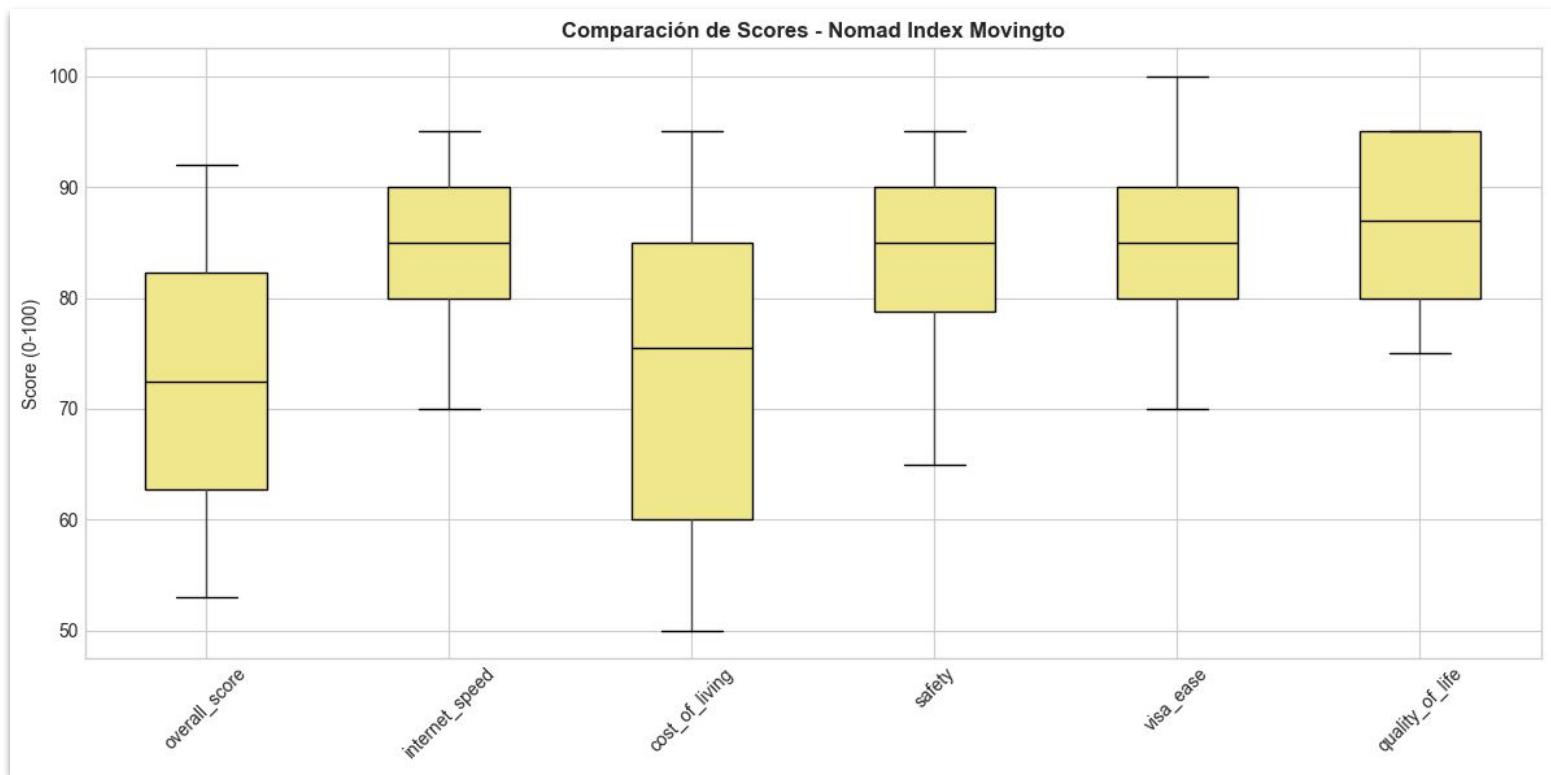
	rank	overall_score	internet_speed	cost_of_living	safety	visa_ease	quality_of_life
count	40.00	40.00	40.0	40.00	40.00	40.00	40.00
mean	20.50	72.50	83.8	73.97	84.25	85.05	87.30
std	11.69	11.69	6.6	13.64	8.33	7.09	6.66
min	1.00	53.00	70.0	50.00	65.00	70.00	75.00
25%	10.75	62.75	80.0	60.00	78.75	80.00	80.00
50%	20.50	72.50	85.0	75.50	85.00	85.00	87.00
75%	30.25	82.25	90.0	85.00	90.00	90.00	95.00
max	40.00	92.00	95.0	95.00	95.00	100.00	95.00

Variables numéricas [Movingto]

Distribución de Scores - Nomad Index Movingto



Variables numéricas [Movingto]





Top 10 países con mejor velocidad de internet para nómadas digitales

Top 10 países con mejor velocidad de internet para nómadas digitales:

country	internet_speed
Estonia	95
Singapore	95
South Korea	95
Netherlands	92
Portugal	90
Germany	90
Japan	90
Taiwan	90
Iceland	90
Norway	90



Top 10 países con peor velocidad de internet para nómadas digitales

Top 10 países con peor velocidad de internet para nómadas digitales:

country	internet_speed
Argentina	70
Colombia	70
Indonesia	75
Uruguay	75
Panama	75
Barbados	75
Montenegro	75
Costa Rica	79
Malaysia	80
Vietnam	80



Top 10 países con mejor seguridad para nómadas digitales

Top 10 países con mejor seguridad para nómadas digitales:

country	safety
Portugal	95
Singapore	95
Japan	95
New Zealand	95
Iceland	95
Norway	95
Finland	95
Spain	92
Estonia	90
Germany	90



Top 10 países con peor seguridad para nómadas digitales

Top 10 países con peor seguridad para nómadas digitales:

country	safety
Colombia	65
Argentina	70
Serbia	70
Mexico	75
Vietnam	75
Panama	75
Montenegro	75
Hungary	75
Romania	75
Bulgaria	75



Top 10 mejores países por facilidad de visa para nómadas digitales

Top 10 mejores países por facilidad de visa para nómadas digitales:

country	visa_ease
Georgia	100
Estonia	98
Portugal	95
Barbados	95
Mexico	94
Spain	92
Thailand	90
Malaysia	90
Costa Rica	90
Malta	90



Top 10 peores países por facilidad de visa para nómadas digitales

Top 10 peores países por facilidad de visa para nómadas digitales:

country	visa_ease
Japan	70
Singapore	75
South Korea	75
Taiwan	75
Australia	75
New Zealand	75
Netherlands	80
Vietnam	80
Canada	80
Hungary	80

Análisis bivariante

Enero 2025



Mergeamos los datasets para análisis bivariante y validar las hipótesis planteadas en el EDA

Cost of Living: variables agregadas del dataset

- `nomad_housing_cost`
- `basic_basket_index`
- `daily_meal_cost`
- `monthly_nomad_cost`
- `local_purchasing_power`
- `housing_salary_ratio`

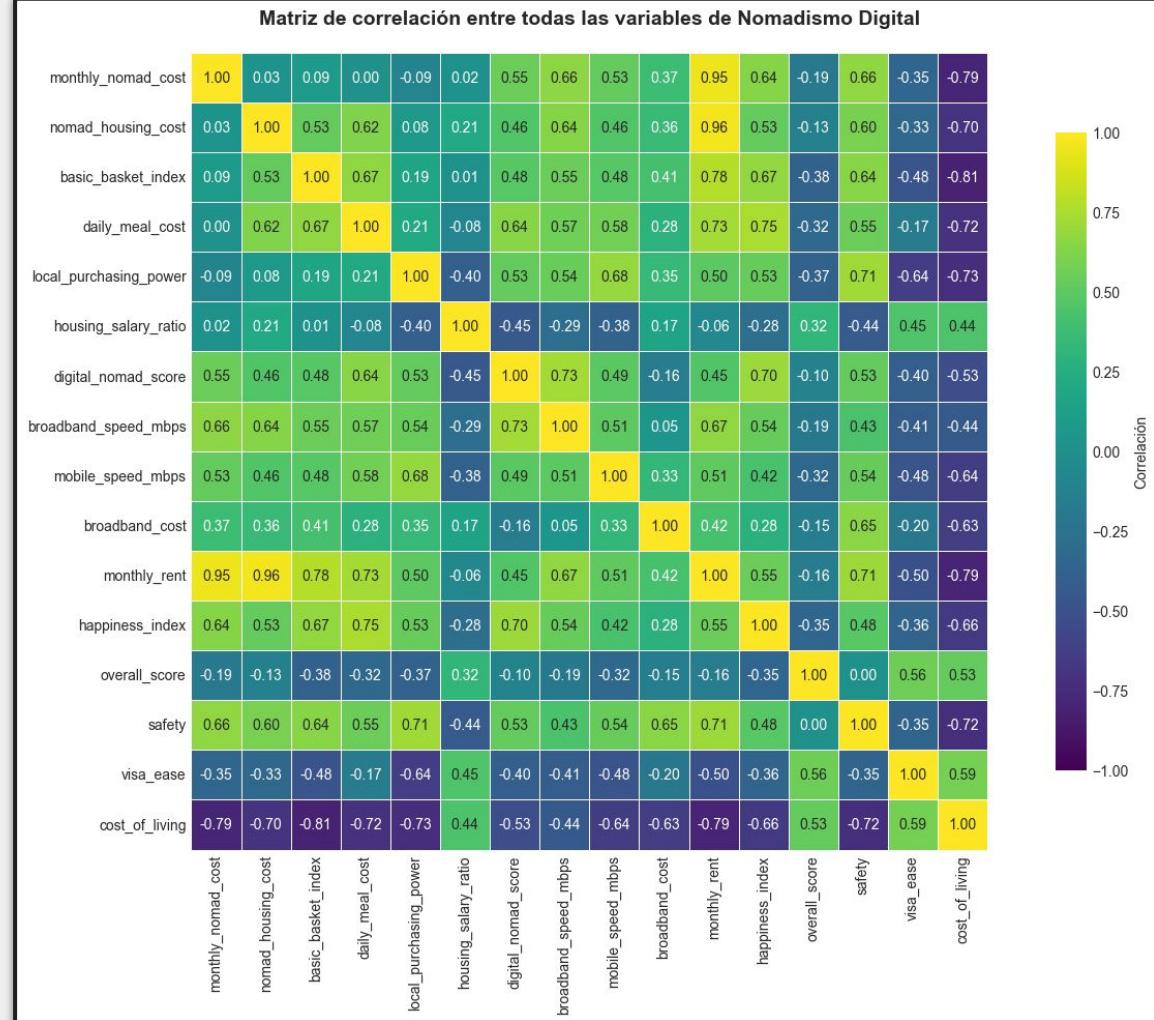
Nomad Index Circleloop:
variables clave del dataset

- `digital_nomad_score`
- `broadband_speed_mbps`
- `happiness_index`
- `monthly_rent`

Nomad Index Movingto:
variables clave del dataset

- `overall_score`
- `safety`
- `visa_ease`
- `cost_of_living`

Matriz de correlación de todas las variables con mapa de calor





¿Cuáles son las variables de cada destino que mejor puntúan para los nómadas digitales?

broadband_speed_mbps	+0.73 (correlación fuerte)
happiness_index	+0.70 (correlación fuerte)
daily_meal_cost	+0.64 (correlación fuerte)
monthly_nomad_cost	+0.55 (correlación fuerte)
cost_of_living	-0.53 (correlación fuerte)
safety	+0.53 (correlación fuerte)
local_purchasing_power	+0.53 (correlación fuerte)
mobile_speed_mbps	+0.49 (correlación moderada)
basic_basket_index	+0.48 (correlación moderada)
nomad_housing_cost	+0.46 (correlación moderada)
monthly_rent	+0.45 (correlación moderada)
housing_salary_ratio	-0.45 (correlación moderada)
visa_ease	-0.40 (correlación moderada)
broadband_cost	-0.16 (correlación débil)
overall_score	-0.10 (correlación débil)

La velocidad de internet es la que tiene un mayor grado de correlación a priori

Después vienen índice de felicidad, coste de comidas fuera de casa y gasto mensual

Validación de hipótesis



Validaciones de hipótesis 1 y 2: gastos mensuales y poder adquisitivo

Hipótesis primarias

- HIPÓTESIS 1

Coste de vida medio: ¿los mejores destinos siempre son los que menos coste mensual exigen?

- HIPÓTESIS 2

Salarios vs coste: ¿la relación entre los salarios y el coste es clave para ser un buen destino?

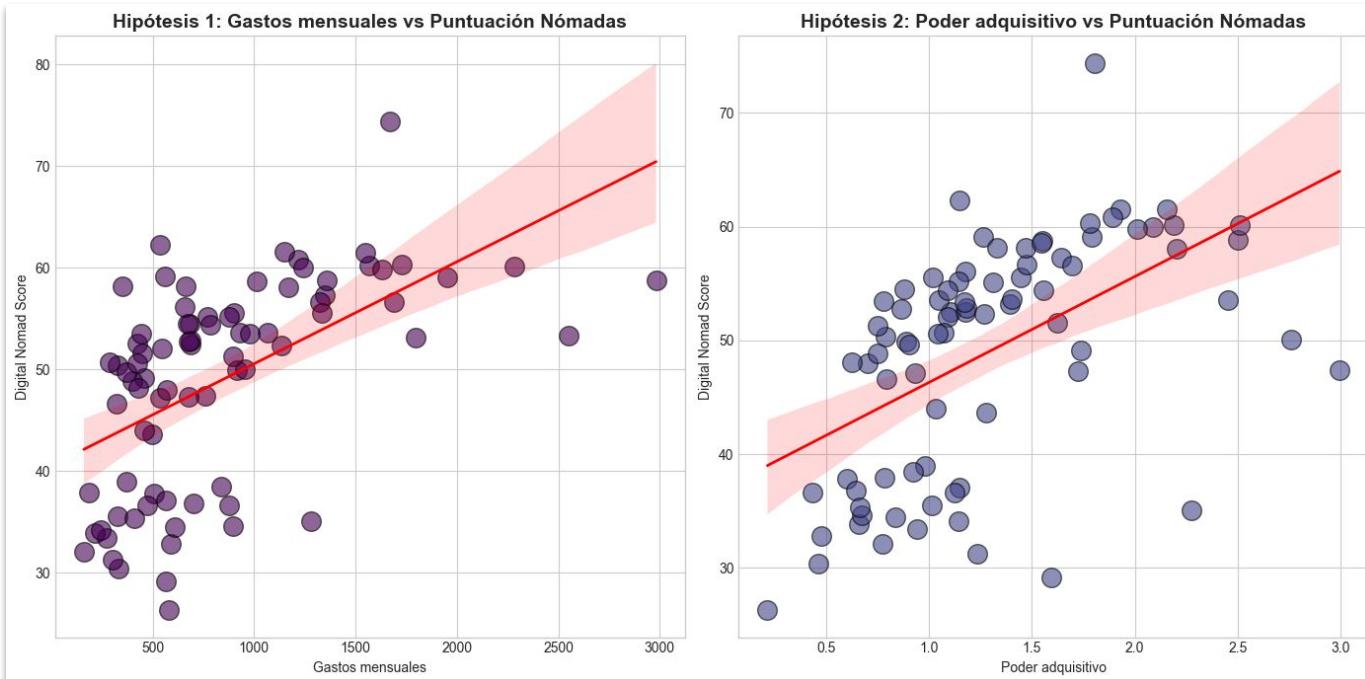


Validaciones de hipótesis 1 y 2: gastos mensuales y poder adquisitivo

Usamos visualizaciones scatterplot para ver la relación entre variables.

Siempre comparamos variable objetivo por la hipótesis con el Digital Nomad Score por país.

Además usamos librería regplot de seaborn para mostrar las líneas de regresión y la correlación.





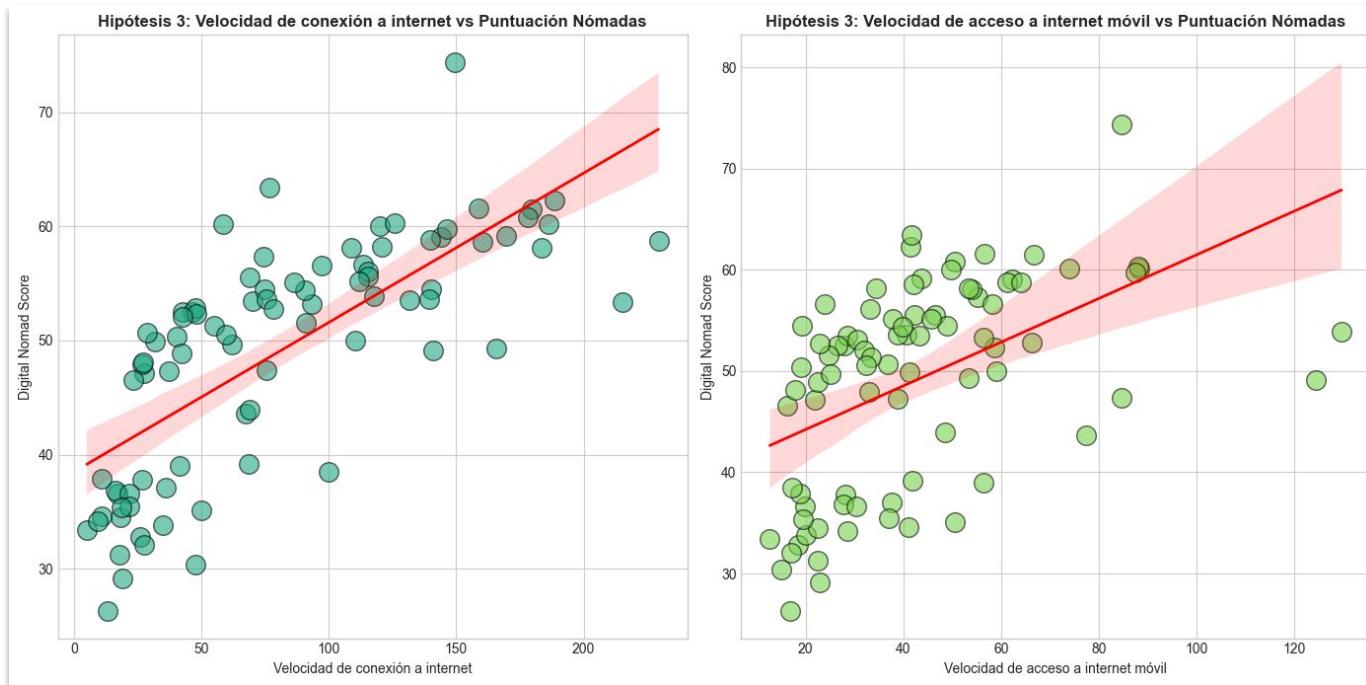
Validaciones de hipótesis 3: velocidad en conexión a internet y coste del servicio

Hipótesis primarias

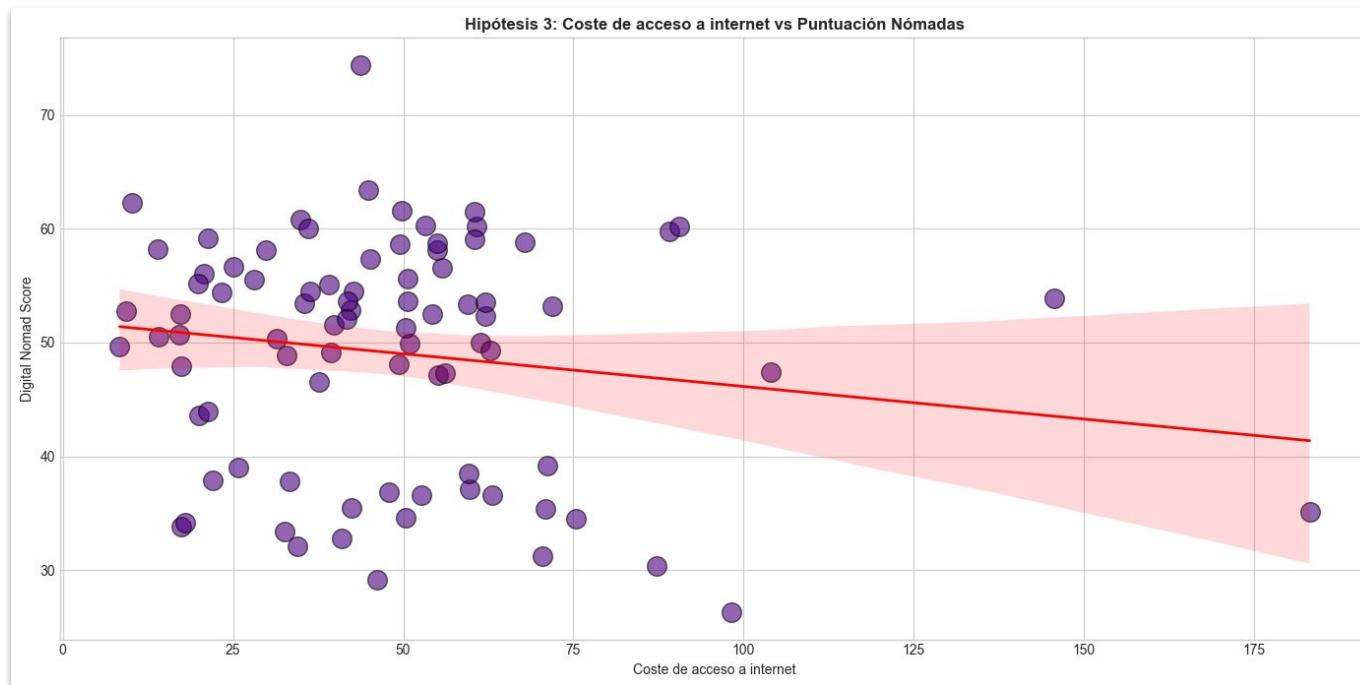
- HIPÓTESIS 3

Conectividad: ¿cuanto más velocidad de conexión a menos precio, mejor destino nómada?

Validaciones de hipótesis 3: velocidad en conexión a internet y coste del servicio



Validaciones de hipótesis 3: velocidad en conexión a internet y coste del servicio





Validaciones de hipótesis 4, 5 y 6: % salario dedicado a vivienda, relación calidad y precio y bienestar social

El análisis se estructura en torno a hipótesis primarias y secundarias:

Hipótesis secundarias

- HIPÓTESIS 4

Vivienda: ¿el alquiler de vivienda es una razón determinante por encima de alimentación?

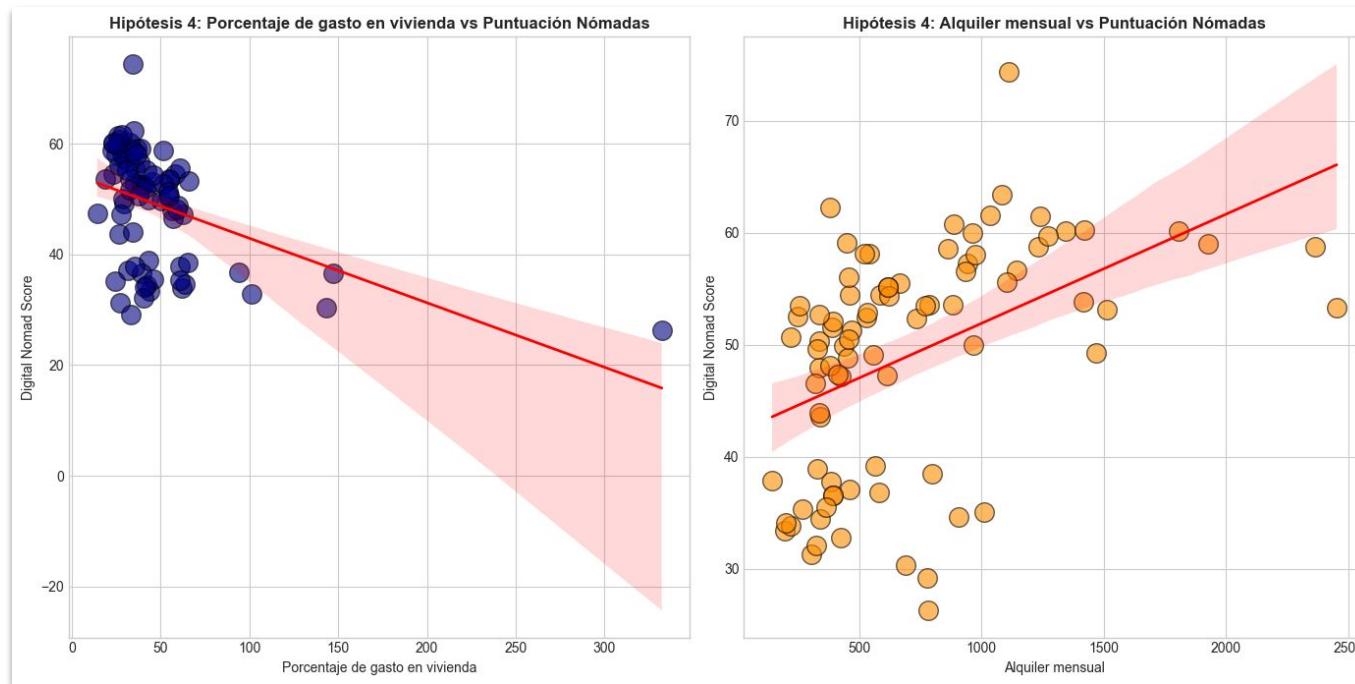
- HIPÓTESIS 5

Más por menos dinero: ¿hay países con mejores condiciones por menos coste mensual?

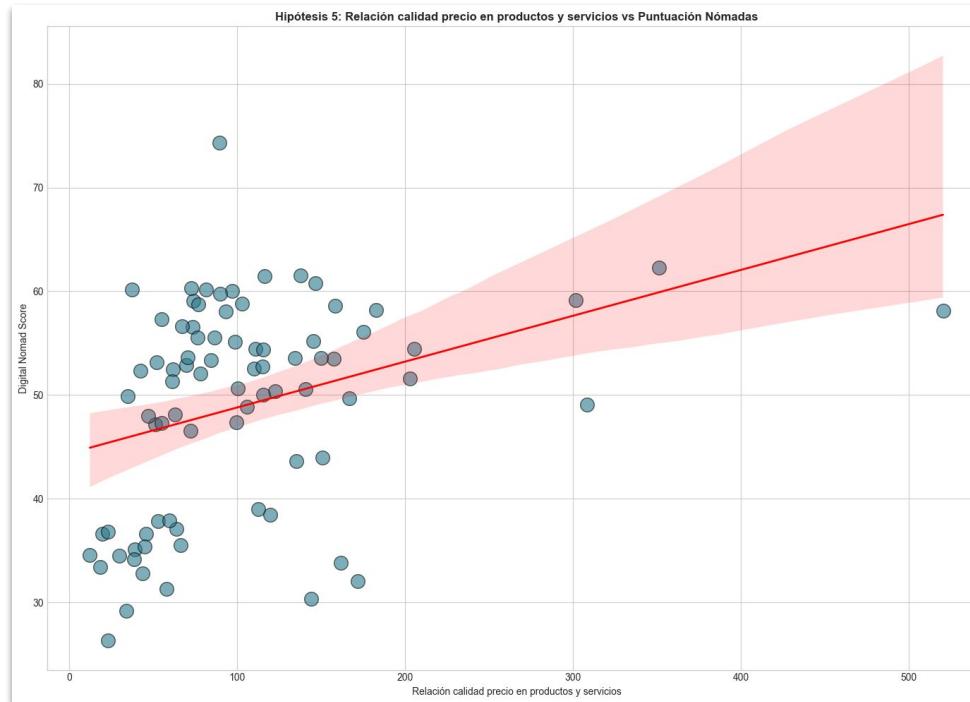
- HIPÓTESIS 6

Bienestar social: ¿los países con buen índice de felicidad atraen más nómadas digitales?

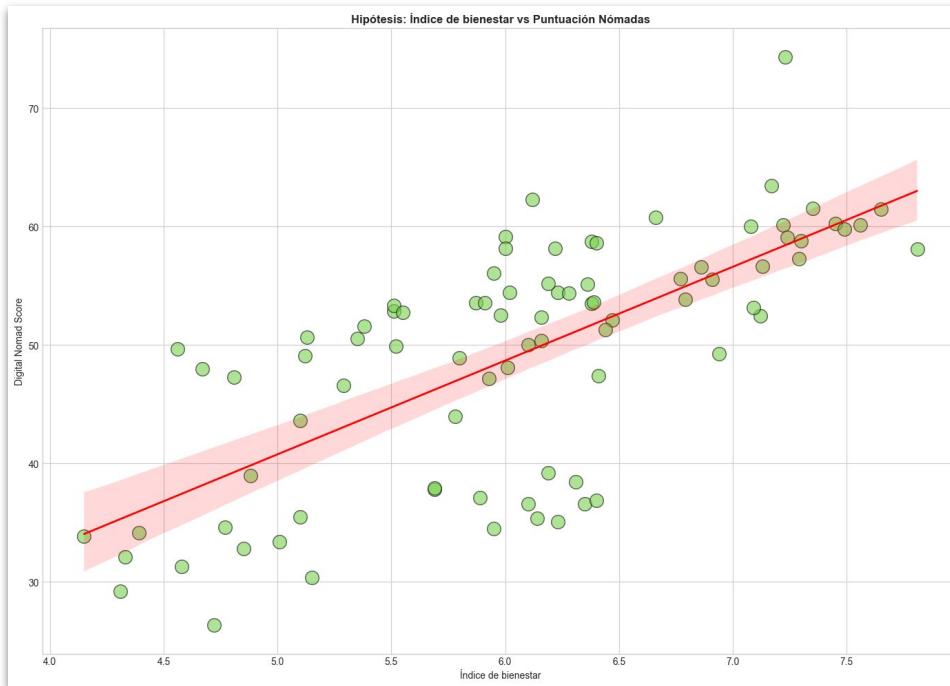
Validaciones de hipótesis 4: porcentaje de salario dedicado a vivienda (alquiler 1 habitación)



Validaciones de hipótesis 5: relación calidad-precio de productos y servicios



Validaciones de hipótesis 6: mayor bienestar social invita a ser un país con nómadas digitales





Validaciones de hipótesis 7, 8 y 9: menor coste vs mayor puntuación, seguridad y facilidad de obtención de visado

El análisis se estructura en torno a hipótesis primarias y secundarias:

Hipótesis secundarias

- HIPÓTESIS 7

Outliers: ¿países fuera de rango por coste y puntuación que son buenos destinos?

- HIPÓTESIS 8

Seguridad: ¿la seguridad es un elemento decisivo en la elección, incluso con costes elevados?

- HIPÓTESIS 9

Facilidad de visado: ¿la accesibilidad legal a los países afecta al ranking de destinos?

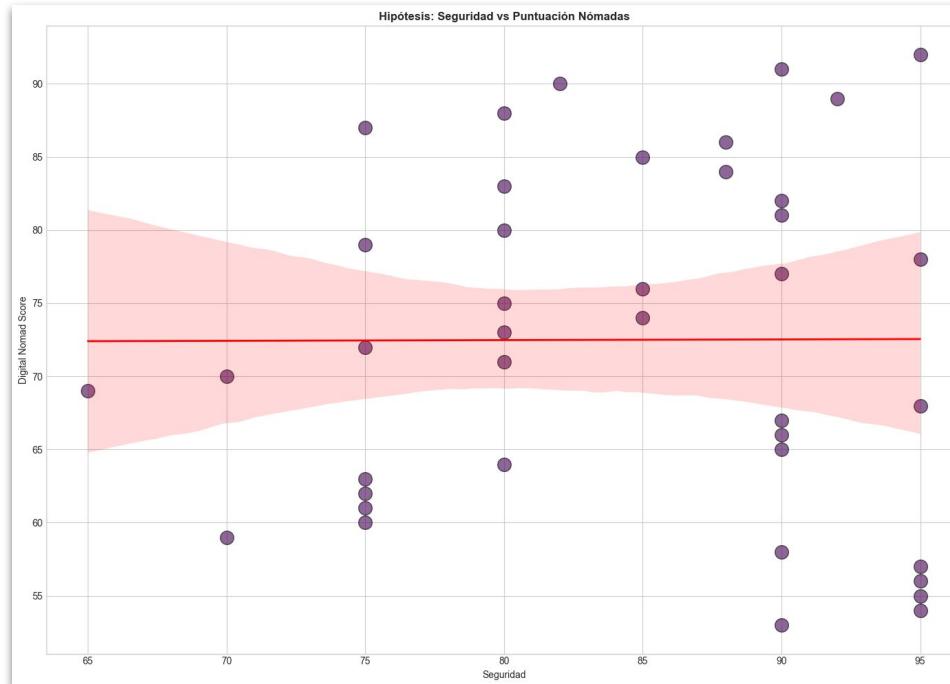


Validaciones de hipótesis 7: estos serían los países con mejor equilibrio gasto / score

Países con menos coste mensual y mejor score en ranking para nómadas digitales

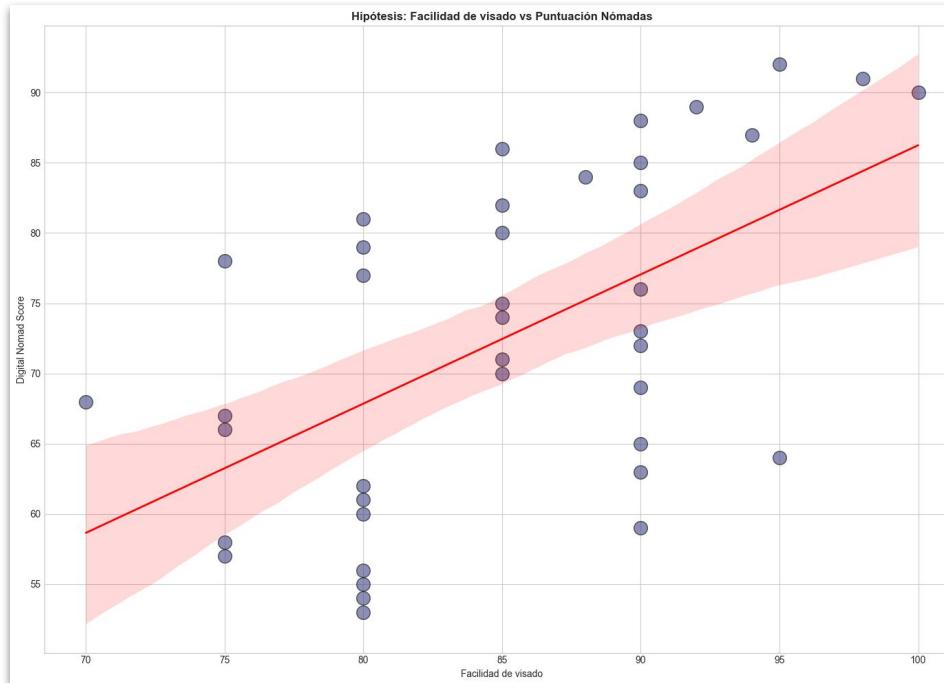
country_name	continent	monthly_nomad_cost	digital_nomad_score
Romania	Europe	536.80000	62.28
Hungary	Europe	562.21750	59.12
Lithuania	Europe	661.18750	58.17
Thailand	Asia	352.85500	58.13
Latvia	Europe	657.97500	56.08
Chile	South America	682.34000	54.46
Estonia	Europe	673.18625	54.45
Brazil	South America	445.74500	53.49
Croatia	Europe	684.62625	52.86
Russia	Europe	678.43500	52.72
Argentina	South America	422.31750	52.53
Costa Rica	North America	688.92000	52.48

Validaciones de hipótesis 8: seguridad como motivo de elección de destino como nómada





Validaciones de hipótesis 9: facilidad de obtención de la visa para elegir destino





Resumen de correlaciones y confirmaciones de hipótesis planteadas en el EDA

broadband_speed_mbps	+0.73 (correlación fuerte)
happiness_index	+0.70 (correlación fuerte)
daily_meal_cost	+0.64 (correlación fuerte)
monthly_nomad_cost	+0.55 (correlación fuerte)
cost_of_living	-0.53 (correlación fuerte)
safety	+0.53 (correlación fuerte)
local_purchasing_power	+0.53 (correlación fuerte)
mobile_speed_mbps	+0.49 (correlación moderada)
basic_basket_index	+0.48 (correlación moderada)
nomad_housing_cost	+0.46 (correlación moderada)
monthly_rent	+0.45 (correlación moderada)
housing_salary_ratio	-0.45 (correlación moderada)
visa_ease	-0.40 (correlación moderada)
broadband_cost	-0.16 (correlación débil)
overall_score	-0.10 (correlación débil)

La velocidad de internet es la que tiene un mayor grado de correlación a priori

Después vienen índice de felicidad, coste de comidas fuera de casa y gasto mensual



Resumen de correlaciones y confirmaciones de hipótesis planteadas en el EDA

Hipótesis primarias

- ✗ **Hipótesis 1:** Existe una relación positiva entre coste y ranking, por lo que un menor gasto mensual no lleva a elegir destino para comenzar una vida como nómada digital.
- ✓ **Hipótesis 2:** A mayor poder adquisitivo, mejor nota al país por parte de los nómadas.
- ✓ **Hipótesis 3:** A mayor velocidad de acceso a internet, mejor el puesto en el ranking.



Resumen de correlaciones y confirmaciones de hipótesis planteadas en el EDA

Hipótesis secundarias

✗ **Hipótesis 4:** El coste en vivienda no es relevante. Comer fuera o el gasto en la cesta sí.

✓ **Hipótesis 5:** Mejor relación calidad precio de productos y servicios correla positivamente.

✓ **Hipótesis 6:** Sí, los nómadas digitales se fijan en el índice de felicidad de los países.



Resumen de correlaciones y confirmaciones de hipótesis planteadas en el EDA

Hipótesis secundarias

- ✓ **Hipótesis 7:** El TOP3 de países con menos coste y mejor puntuación son Rumanía, Hungría y Lituania, tres países del este de Europa.
- ✗ **Hipótesis 8:** La seguridad en el país de destino no es un elemento clave para el ranking.
- ✓ **Hipótesis 9:** La facilidad para conseguir la visa, en cambio, sí lo es.

Gracias :_]