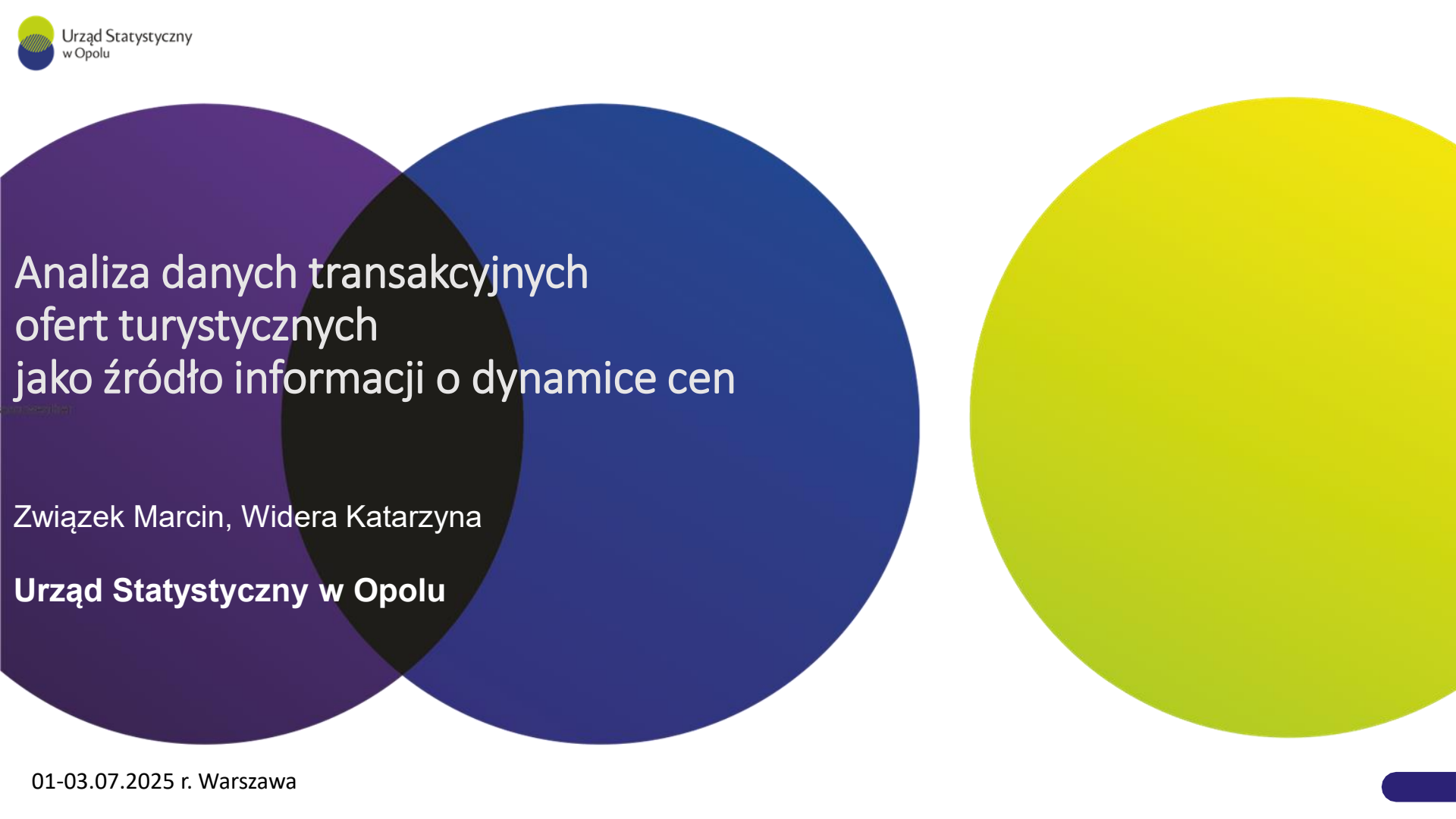




Analiza danych transakcyjnych ofert turystycznych jako źródło informacji o dynamice cen

Związek Marcin, Widera Katarzyna

Urząd Statystyczny w Opolu

Plan prezentacji

1.

Walidacja zbioru danych

2.

Przetwarzanie i analiza danych

3.

Budowa modelu

4.

Problemy i wyzwania

Źródło danych A

Miesiąc	Kraj	Region	Standard	Wyżywienie	Port wylotowy	Liczba Noc	Średnia cena
2025-01-01 00:00:00.000	Hiszpania	Wyspy Kanaryjskie	4	A	WAW	7	3784,06449165402
2025-01-01 00:00:00.000	Hiszpania	Wyspy Kanaryjskie	4	A	KTW	7	3790,13840304182
2025-01-01 00:00:00.000	Egipt	Marsa Alam	5	A	KTW	7	2900,65352112676
2025-01-01 00:00:00.000	Egipt	Marsa Alam	5	A	WAW	7	2796,72740524781

Źródło danych B

Miesiąc wylotu	Lotnisko wylotowe	Lotnisko docelowe	Kraj	Liczba gwiazdek	Typ wyżywienia	Długość pobytu	Liczba Pax	Cena
5	BZG	AYT	Turcja	5	A	4	2	3804
5	BZG	AYT	Turcja	5	A	7	430	1306636
5	BZG	AYT	Turcja	5	A	8	10	40925
5	BZG	AYT	Turcja	5	A	10	10	47718
5	BZG	AYT	Turcja	5	A	11	4	22464
5	BZG	AYT	Turcja	5	A	14	10	60734
5	BZG	AYT	Turcja	5	X	7	4	8396
5	BZG	AYT	Turcja	3	A	7	18	41056
5	BZG	AYT	Turcja	3	H	7	2	4432
5	BZG	AYT	Turcja	3	U	7	2	5092

Walidacja danych cd

Źródło A

Porównanie liczby kategorii

Styczeń		Maj	
Miesiąc	1	Miesiąc	1
Kraj	14	Kraj	9
Region	18	Region	28
Standard	3	Standard	3
Wyżywienie	3	Wyżywienie	3
Port wylotowy	6	Port wylotowy	5
Liczba nocy	5	Liczba nocy	5

Źródło B

Porównanie liczby kategorii

Styczeń		Maj	
Miesiąc wylotu	1	Miesiąc wylotu	1
Lotnisko wylotowe	13	Lotnisko wylotowe	19
Lotnisko docelowe	53	Lotnisko docelowe	85
Kraj	31	Kraj	35
Liczba gwiazdek	8	Liczba gwiazdek	11
Typ wyżywienia	8	Typ wyżywienia	8
Liczba Pax	148	Długość pobytu	21
		Liczba Pax	285

Cel

- określenie ceny dla reprezentanta (standaryzacja informacji),
- identyfikacja zmiennych określających reprezentanta.

Cena jednostkowa obliczana według wzoru:

$$PPD_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\text{cena}_{i,t}}{\frac{\text{liczba podróży}_{i,t}}{\text{liczba dni}_{i,t}}}$$

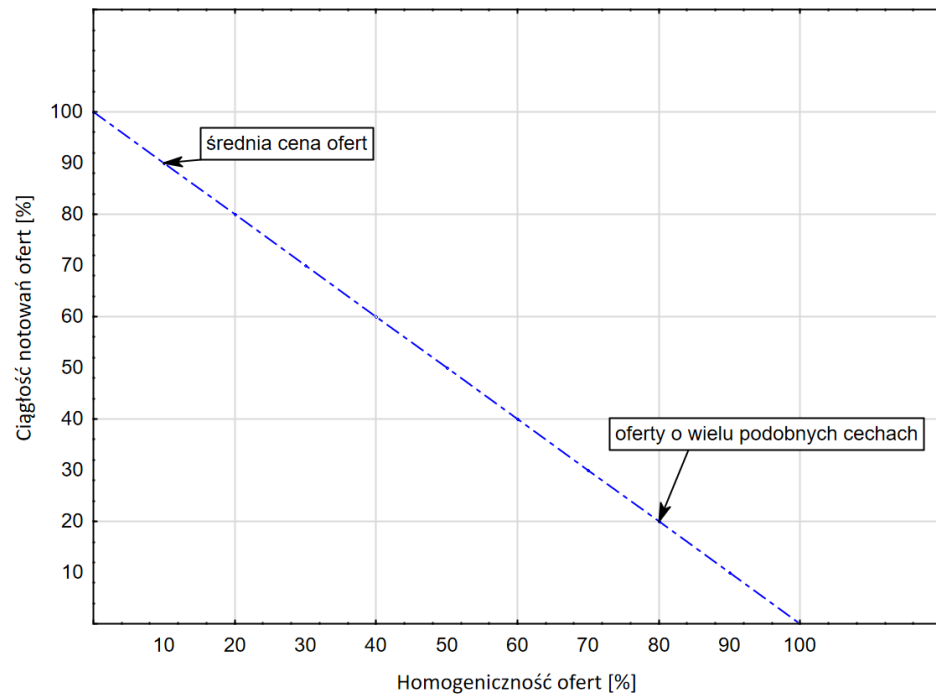
i – numer transakcji w okresie t ,

N_t -liczba transakcji w okresie t .

Potencjalne zmienne określające reprezentanta:

Kraj	Region	Standard
Liczba nocy	Lotnisko wylotowe	Wyżywienie

Ciągłość a homogeniczność



Źródło: EUROSTAT REVIEW ON NATIONAL ACCOUNTS AND MACROECONOMIC INDICATORS 2/2019 PDF ISSN 1977-978X

Reprezentant

destynacja

standard

wyżywienie

Wagi do wskaźnika cen

Dane transakcyjne.

Informacje o % udziale reprezentanta w przychodzie biura podróży.

Wagi do wskaźnika cen.

WYBRANE DESTYNACJE

Egipt	Grecja	Hiszpania	Dominikana	Portugalia
Republika Zielonego Przylądka	Tanzania	Tunezja	Turcja	Włochy

Model regresji dla reprezentantów

OLS Regression Results						
=====						
Dep. Variable:	y	R-squared:	0.927			
Model:	OLS	Adj. R-squared:	0.916			
Method:	Least Squares	F-statistic:	83.63			
Date:	Wed, 18 Jun 2025	Prob (F-statistic):	9.90e-40			
Time:	10:18:02	Log-Likelihood:	-518.13			
No. Observations:	92	AIC:	1062.			
Df Residuals:	79	BIC:	1095.			
Df Model:	12					
Covariance Type:	nonrobust					
=====						
		coef	std err	t	P> t	[0.025 0.975]

const		1146.7008	45.624	25.134	0.000	1055.889 1237.513
okres_2025-02-01		127.6588	19.573	6.522	0.000	88.700 166.618
Kraj_Egipt		-856.8370	48.409	-17.700	0.000	-953.192 -760.482
Kraj_Grecja		-873.7232	54.690	-15.976	0.000	-982.580 -764.866
Kraj_Hiszpania		-693.6738	48.480	-14.308	0.000	-790.171 -597.177
Kraj_Portugalia		-826.4725	47.184	-17.516	0.000	-920.389 -732.556
Kraj_Republika Zielonego Przylądka		-511.3703	49.499	-10.331	0.000	-609.896 -412.844
Kraj_Tanzania		-312.9834	52.210	-5.995	0.000	-416.906 -209.061
Kraj_Tunezja		-927.4682	49.499	-18.737	0.000	-1025.994 -828.942
Kraj_Turcja		-899.2394	60.115	-14.959	0.000	-1018.895 -779.584
Kraj_Włochy		-908.0991	72.117	-12.592	0.000	-1051.644 -764.554
Wyżywienie_H		83.9400	21.421	3.919	0.000	41.302 126.578
Standard_5		152.2663	16.567	9.191	0.000	119.291 185.242
=====						
Omnibus:	0.859	Durbin-Watson:	1.451			
Prob(Omnibus):	0.651	Jarque-Bera (JB):	0.520			
Skew:	0.174	Prob(JB):	0.771			
Kurtosis:	3.122	Cond. No.	22.9			
=====						

Indeksy cen i indeksy hedoniczne cen dla reprezentantów

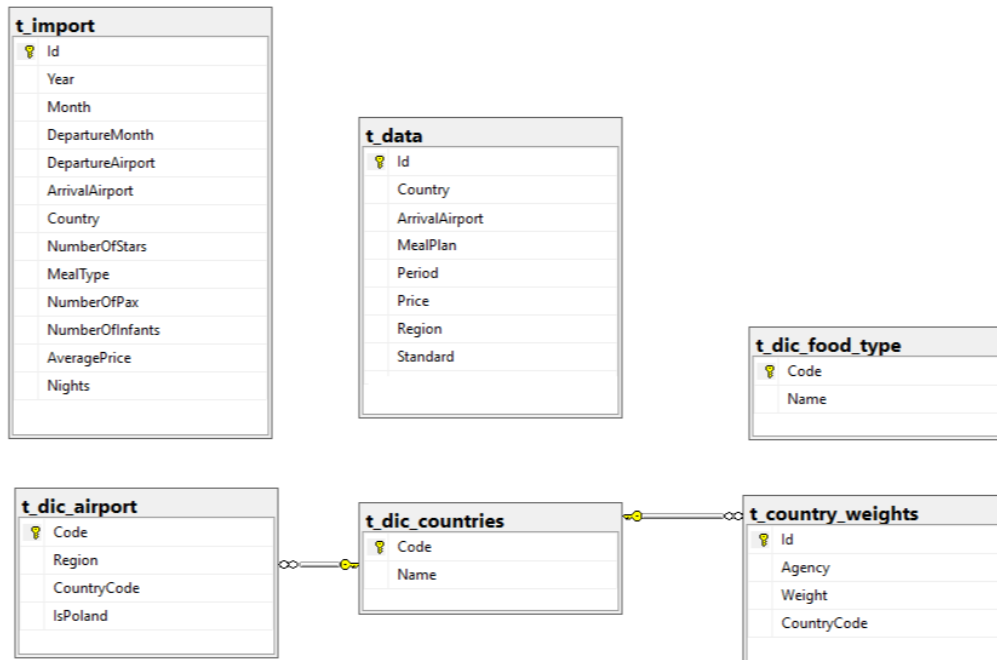
okres	ceny średnie	ceny hedoniczne	I	I hedoniczny	zmiana %	I (styczeń =100)	I hedoniczny (styczeń = 100)
styczeń	554,97	554,03				100,0%	100,0%
luty	676,04	671,50	121,8%	121,0%	-0,8%	121,8%	121,2%
marzec	519,04	516,69	76,8%	76,4%	-0,3%	93,5%	76,9%
kwiecień	431,04	442,56	83,0%	85,3%	2,2%	77,7%	85,7%
maj	403,77		93,7%			72,8%	

Średniookresowe tempo zmian 92,4%

Wyzwania na dalsze prace

Kraj	Standard	Wyżywienie	202501	202502	202503	202504	202505
Hiszpania	4	A	509,777503379921	615,977728979511	532,444215164965	543,883343665301	522,97151453687
Hiszpania	4	H	517,173707157603	625,877769758288	497,909663839475	456,861731290807	473,225197421626
Hiszpania	5	H	821,597802197802	1015,28761904762	NULL	NULL	803,60599078341
Portugalia	3	F	238,95619047619	414,100460308487	374,168087730588	368,010966064859	445,458928571429
Portugalia	3	H	323,332365704903	489,732046983543	408,979127096383	431,883345178867	409,499779541446
Portugalia	4	H	349,684534420751	519,94416696256	425,478212597096	442,130740817962	514,064940587435
Portugalia	5	A	440,79476861167	587,316326530612	437,589981447124	419,855728429986	NULL
Portugalia	5	F	394,330952380952	NULL	512,760504201681	NULL	NULL
Portugalia	5	H	478,103040540541	628,546356776845	555,644287065394	658,267208767209	654,176452567071
Republika...	4	A	555,620554619307	801,321428571429	621,207263922518	571,242920075519	580,684582743989

BUDOWA BAZY DANYCH



Wyzwania na dalsze prace cd

- ☐ Ciągła walidacja danych w bazie zgodna z ustaleniami.
- ☐ Modyfikacja słowników nazw zmiennych i ich kategorii.
- ☐ Budowa modeli hedonicznych, obliczanie korekt jakościowych, wyznaczenie cen hedonicznych dla kolejnych okresów.
- ☐ Obliczanie wskaźnika cen ofert turystycznych.
- ☐ Obliczanie skorygowanego wskaźnika cen ofert turystycznych.

- EURONA — Eurostat Review on National Accounts and Macroeconomic Indicators — Issue No 1 (2019) <https://ec.europa.eu/eurostat/web/products-statistical-books/-/ks-gp-19-001>
- Technical Paper Forecasting HICP package holidays with forward-looking booking data (2024)
BundesBank Forecasting HICP package holidays with forward-looking booking data
- Statistics Paper Series Chain linking over December and methodological changes in the HICP: view from a central bank perspective No 40 (2021) European Central Bank
- EURONA — Eurostat Review on National Accounts and Macroeconomic Indicators — Issue 2020 <https://ec.europa.eu/eurostat/web/products-statistical-books/-/ks-gp-20-001>
- Białek J., Kłopotek M., Panek T. (2022): *Nowoczesne technologie i nowe źródła danych w pomiarze inflacji*. GUS, Biblioteka Wiadomości Statystycznych

Dziękujemy za uwagę