

Научная работа на тему: «Эконометрическое исследование зависимости стоимости туристической путёвки от различных факторов»

ВЫПОЛНИЛИ:

СТУДЕНТКИ ГРУППЫ ЭК-21

БОРОВСКАЯ Л.К.

КИСЕЛЕВА М.А.

РУКОВОДИТЕЛИ:

К.Э.Н., ПРОФ. ДОМОЖИРОВА
О.В.

СТ. ПРЕП. ХЛЕБЕНСКИХ Л.В.

Существуют факторы, отвечающие за составляющие элементы туристической путевки, которые напрямую влияют на цену. Мы выбрали 9 основных факторов:

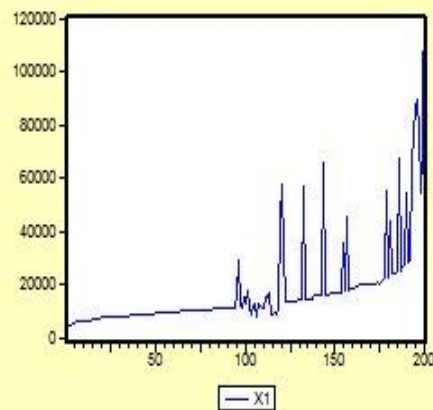
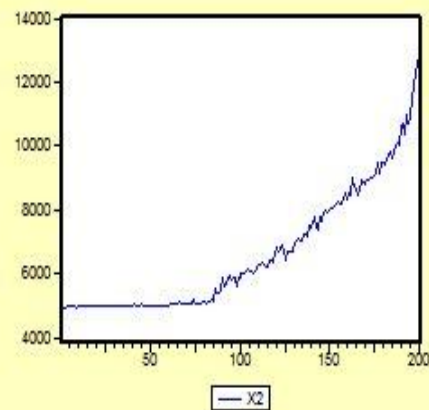
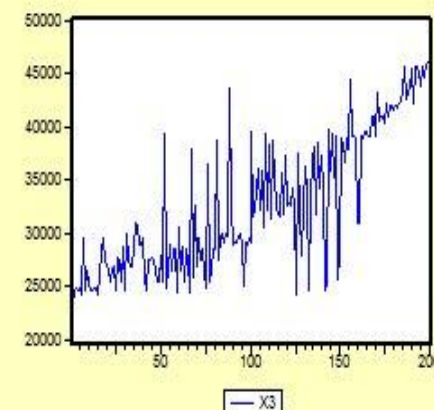
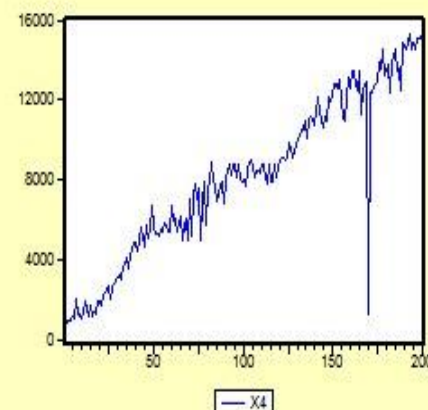
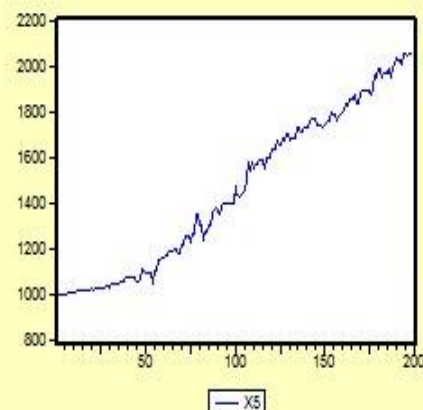
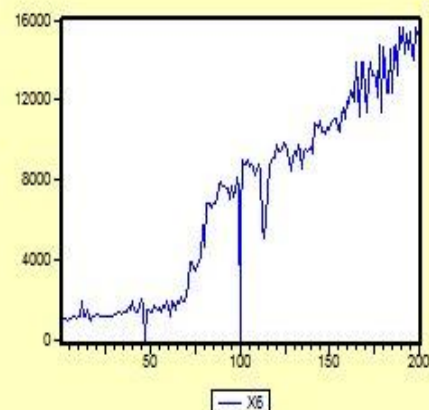
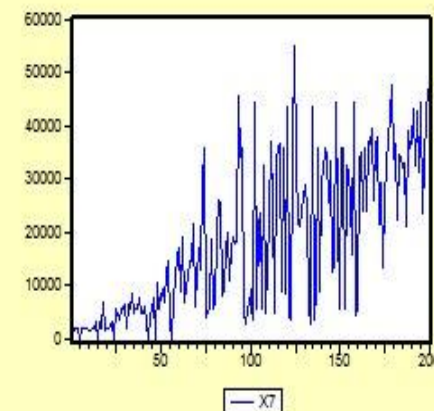
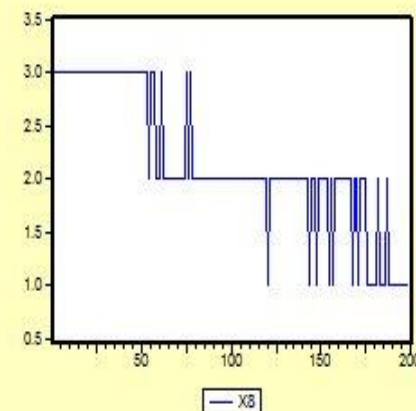
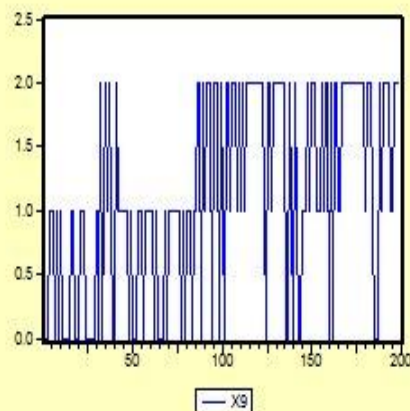
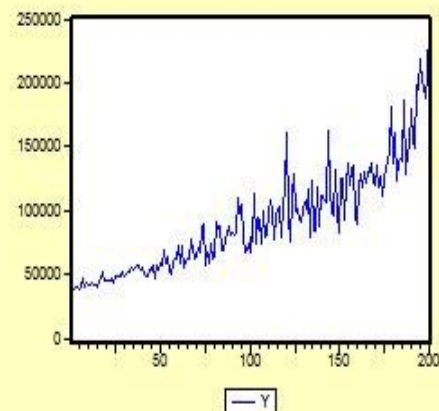
- стоимость номера (X1);
- стоимость питания (X2);
- стоимость переезда (X3);
- стоимость трансфера (X4);
- стоимость страховки (X5);
- услуги гида – переводчика (X6);
- стоимость экскурсий (X7);
- тип номера (X8);
- сезон (X9).

Анализ влияния различных факторов на формирование цены туристической путевки в Индию

Результирующий признак в представленном исследовании – стоимость туристической путевки в Индию. Исходные данные эконометрического исследования представлены в таблице.

Представим
исходные
данные в виде
линейных
графиков

Уже на этапе
построения
модели можно
заметить, что
динамика
нескольких
показателей
совпадает



Следующий шаг в исследование выбранных случайных переменных – это определение тесноты связи между ними. Для этого необходимо рассчитать линейный коэффициент корреляции между X и Y.

	Y	X9	X8	X7	X6	X5	X4	X3	X2	X1
Y	1.000000	0.462115	-0.864233	0.832305	0.893247	0.905100	0.876712	0.840915	0.921608	0.841248
X9	0.462115	1.000000	-0.441718	0.448984	0.507779	0.508581	0.487856	0.387015	0.436708	0.281846
X8	-0.864233	-0.441718	1.000000	-0.735167	-0.835394	-0.851507	-0.843289	-0.755388	-0.778070	-0.662849
X7	0.832305	0.448984	-0.735167	1.000000	0.727475	0.739309	0.717593	0.615123	0.684625	0.483000
X6	0.893247	0.507779	-0.835394	0.727475	1.000000	0.970661	0.924660	0.821677	0.907437	0.630296
X5	0.905100	0.508581	-0.851507	0.739309	0.970661	1.000000	0.941991	0.835258	0.926422	0.645738
X4	0.876712	0.487856	-0.843289	0.717593	0.924660	0.941991	1.000000	0.789708	0.870220	0.618459
X3	0.840915	0.387015	-0.755388	0.615123	0.821677	0.835258	0.789708	1.000000	0.845984	0.622249
X2	0.921608	0.436708	-0.778070	0.684625	0.907437	0.926422	0.870220	0.845984	1.000000	0.759796
X1	0.841248	0.281846	-0.662849	0.483000	0.630296	0.645738	0.618459	0.622249	0.759796	1.000000

Стандартный вид уравнения:

$$y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + b_4 \cdot x_4 + b_5 \cdot x_5 + b_6 \cdot x_6 + b_7 \cdot x_7 + b_8 \cdot x_8 + b_9 \cdot x_9$$

Dependent Variable: Y				
Method: Least Squares				
Date: 05/25/19 Time: 00:16				
Sample: 1 200				
Included observations: 200				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.60E-10	3.32E-10	-1.685724	0.0935
X1	1.000000	1.87E-15	5.35E+14	0.0000
X2	1.000000	3.33E-14	3.00E+13	0.0000
X3	1.000000	5.43E-15	1.84E+14	0.0000
X4	1.000000	1.35E-14	7.42E+13	0.0000
X5	1.000000	2.79E-13	3.58E+12	0.0000
X6	1.000000	1.57E-14	6.38E+13	0.0000
X7	1.000000	1.96E-15	5.11E+14	0.0000
X8	4.57E-10	6.24E-11	7.315769	0.0000
X9	0.000000	2.65E-11	0.000000	1.0000
R-squared	1.000000	Mean dependent var	91155.28	
Adjusted R-squared	1.000000	S.D. dependent var	41485.90	
S.E. of regression	2.52E-10	Akaike info criterion	-41.31563	
Sum squared resid	1.21E-17	Schwarz criterion	-41.15071	
Log likelihood	4141.563	F-statistic	5.99E+29	
Durbin-Watson stat	1.266563	Prob(F-statistic)	0.000000	

Запишем полученное уравнение множественной регрессии:

$$y = (-5,60E-10) + 1,00 \cdot x_1 + 1,00 \cdot x_2 + 1,00 \cdot x_3 + 1,00 \cdot x_4 + 1,00 \cdot x_5 + 1,00 \cdot x_6 + 1,00 \cdot x_7 + (4,54E-10) \cdot x_8 + 0,00 \cdot x_9$$

Затем рассмотрим значение F-критерий Фишера. С помощью этого критерия оценивают качество регрессионной модели в целом и по параметрам, значимость $F = 3,43457E+30$, то есть уравнение регрессии значимо.

Применим тест Уайта и проанализируем модель на гетероскедастичность:

По тесту Уайта уравнение в целом значимо, $F_{набл} = 7,88 > F_{табл}$. По t-статистике значимость коэффициентов при x незначительна. Следовательно, необходимо признать наличие гетероскедастичности по тесту Уайта.

White Heteroskedasticity Test:				
F-statistic	7.877342	Probability	0.000000	
Obs*R-squared	87.85353	Probability	0.000000	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 05/25/19 Time: 20:54				
Sample: 1 200				
Included observations: 200				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.66E-18	6.17E-19	-2.689374	0.0078
X1	1.38E-25	2.35E-24	0.058822	0.9532
X1^2	-1.02E-29	2.37E-29	-0.431554	0.6666
X2	8.32E-23	8.28E-23	1.004666	0.3164
X2^2	-3.49E-27	4.13E-27	-0.845697	0.3988
X3	-7.77E-23	2.06E-23	-3.769602	0.0002
X3^2	1.26E-27	3.14E-28	4.023834	0.0001
X4	-1.02E-23	1.54E-23	-0.661860	0.5089
X4^2	6.30E-28	1.10E-27	0.571245	0.5685
X5	3.95E-21	8.61E-22	4.587659	0.0000
X5^2	-1.32E-24	2.98E-25	-4.446820	0.0000
X6	-9.76E-23	1.32E-23	-7.425004	0.0000
X6^2	4.52E-27	8.91E-28	5.071441	0.0000
X7	1.17E-24	2.37E-24	0.494458	0.6216
X7^2	-3.42E-29	4.92E-29	-0.695551	0.4876
X8	2.02E-19	1.09E-19	1.854852	0.0652
X8^2	-6.97E-20	2.67E-20	-2.604867	0.0100
X9	-1.39E-20	3.12E-20	-0.445898	0.6562
X9^2	3.47E-21	1.49E-20	0.232702	0.8163
R-squared	0.439268	Mean dependent var	6.04E-20	
Adjusted R-squared	0.383504	S.D. dependent var	1.19E-19	
S.E. of regression	9.33E-20	Sum squared resid	1.57E-36	
F-statistic	7.877342	Durbin-Watson stat	2.109002	
Prob(F-statistic)	0.000000			

Проведем тест Парка также для выявления гетероскедастичности.

Dependent Variable: LOG(RESID01^2)				
Method: Least Squares				
Date: 05/26/19 Time: 12:58				
Sample(adjusted): 4 200				
Included observations: 144				
Excluded observations: 53 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-28.02544	21.51372	-1.302678	0.1949
LOG(X1)	-0.441554	0.701233	-0.629682	0.5300
LOG(X2)	-1.343974	3.049432	-0.440729	0.6601
LOG(X3)	1.140228	2.074531	0.549632	0.5835
LOG(X4)	1.055690	0.660194	1.599061	0.1122
LOG(X5)	-3.811094	5.334936	-0.714366	0.4762
LOG(X6)	0.390297	0.899651	0.433831	0.6651
LOG(X7)	0.240424	0.322594	0.745284	0.4574
LOG(X8)	-1.047039	1.163077	-0.900232	0.3696
LOG(X9)	-0.393672	0.691847	-0.569016	0.5703
R-squared	0.068863	Mean dependent var	-45.75582	
Adjusted R-squared	0.006324	S.D. dependent var	2.426259	
S.E. of regression	2.418575	Akaike info criterion	4.671150	
Sum squared resid	783.8339	Schwarz criterion	4.877387	
Log likelihood	-326.3228	F-statistic	1.101116	
Durbin-Watson stat	2.063539	Prob(F-statistic)	0.366311	

Применим указанные преобразования, используя все 9 весов и для полученных моделей проведем тест Уайта и выберем наилучшую модель.

Вес	X1	X2	X3	X4	X5
Адекватность	R²=0,71	R²=0,84	R²=0,71	R²=0,77	R²=0,82
	F=479,38	F=1116,48	F=478,16	F=657,76	F=897,24
Тест Уайта	R²=0,07	R²=0,04	R²=0,09	R²=0,13	R²=0,20
	F=7,32	F=3,75	F=9,98	F=14,68	F=24,99
Вес	X6	X7	X8	X9	
Адекватность	R²=0,78	R²=0,69	R²=0,75	R²=0,21	
	F=781,66	F=446,43	F=584,40	F=53,77	
Тест Уайта	R²=0,19	R²=0,11	R²=0,14	R²=0,003	
	F=23,64	F=12,21	F=16,7	F=0,27	

ВЫВОДЫ

- ▶ Проведенные расчеты позволяют в качестве преобразования выбрать вес x_2 . Проведя качественный и количественный анализ данных, можно сделать вывод о том, что на стоимость туристической путевки влияет множество факторов. Модель является значимой, а факторы оказывают сильное влияние, особенно в совокупности.
- ▶ Из всех факторов наибольшее влияние оказывает фактор стоимость питания (x_2) исходя из преобразований с помощью метода наименьших квадратов. Следовательно, чтобы изменить стоимость туристической путевки, нужно изменить фактор стоимость питания.