

Приложение к задаче 1.

Таблица 1. Описательные статистики

Variable	Obs	Mean	Std. dev.	Min	Max
exam_score	6,607	67.23566	3.890456	55	100
male	6,607	.5772665	.4940312	0	1
hours_stud~d	6,607	19.97533	5.990594	1	44
male_hours~s	6,607	11.51854	10.85631	0	39
previous_s~s	6,607	75.07053	14.39978	50	100
attendance	6,607	79.97745	11.54747	60	100
sleep_hours	6,607	7.02906	1.46812	4	10
physical_a~y	6,607	2.96761	1.031231	0	6

Таблица 2. Корреляционная матрица

	hours_~d	male	previo~s	attend~e	sleep_~s	physic~y
hours_stud~d	1.0000					
male	-0.0042	1.0000				
previous_s~s	0.0248	0.0010	1.0000			
attendance	-0.0099	0.0080	-0.0202	1.0000		
sleep_hours	0.0110	-0.0075	-0.0218	-0.0159	1.0000	
physical_a~y	0.0046	0.0052	-0.0113	-0.0224	-0.0004	1.0000

Регрессия 1. Оценка модели по полной выборке

Source	SS	df	MS	Number of obs	=	6,607
Model	57382.8232	5	11476.5646	F(5, 6601)	=	1778.19
Residual	42603.2556	6,601	6.45406083	Prob > F	=	0.0000
Total	99986.0787	6,606	15.1356462	R-squared	=	0.5739
				Adj R-squared	=	0.5736
				Root MSE	=	2.5405

exam_score	Coefficient	Std. err.	t	P> t	[95% conf. interval]
hours_studied	.2901824	.0052199	55.59	0.000	.2799497 .3004151
previous_scores	.0475931	.0021725	21.91	0.000	.0433343 .0518518
attendance	.1987214	.0027085	73.37	0.000	.1934117 .204031
sleep_hours	-.0230293	.0212998	-1.08	0.280	-.0647838 .0187251
physical_activity	.1545782	.0303204	5.10	0.000	.0951404 .2140161
_cons	41.67625	.3462125	120.38	0.000	40.99756 42.35494

Таблица 3. VIF (регрессия 1)

Variable	VIF	1/VIF
previous_s~s	1.00	0.998345
attendance	1.00	0.998728
sleep_hours	1.00	0.999129
hours_stud~d	1.00	0.999143
physical_a~y	1.00	0.999336
Mean VIF	1.00	

Таблица 4. Тест Уайта (регрессия 1)

chi2(20) = 27.30
Prob > chi2 = 0.1270

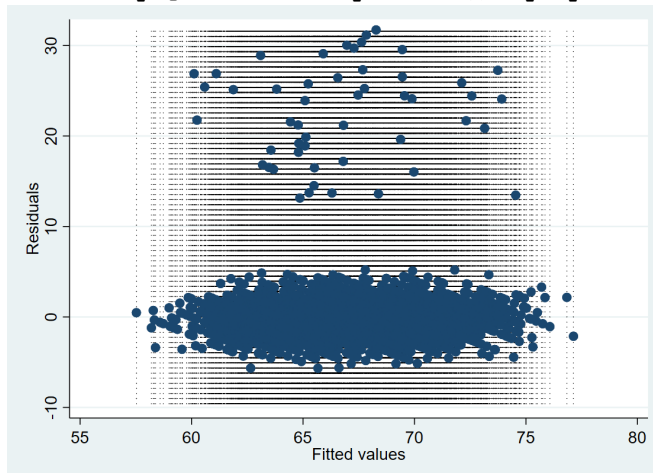
Таблица 5. Тест Рамсея (регрессия 1)

F(3, 6598) = 2.23
Prob > F = 0.0821

Таблица 6. Ковариационная матрица коэффициентов регрессии 5

	hours_stud~d	previous_s~s	attendance	attendance2	sleep_hours	physical_a~y	_cons
hours_stud~d	.00002726						
previous_s~s	-2.815e-07	4.720e-06					
attendance	-3.290e-06	-6.377e-07	.00175727				
attendance2	2.135e-08	4.739e-09	-.00001093	6.826e-08			
sleep_hours	-1.266e-06	1.035e-06	1.689e-06	-4.685e-09	.00045371		
physical_a~y	-7.238e-07	7.863e-07	-8.942e-06	6.746e-08	6.724e-07	.00091945	
_cons	-.00038859	-.00033827	-.06904544	.00042743	-.00334802	-.00250323	2.7961701

Рис. 1. График остатки-прогнозы для регрессии 1



Регрессия 2. Только девушки

Source	SS	df	MS	Number of obs	=	2,793
Model	25694.6417	5	5138.92833	F(5, 2787)	=	708.95
Residual	20201.8481	2,787	7.24859998	Prob > F	=	0.0000
				R-squared	=	0.5598
				Adj R-squared	=	0.5590
Total	45896.4898	2,792	16.4385708	Root MSE	=	2.6923

exam_score	Coefficient	Std. err.	t	P> t	[95% conf. interval]
hours_studied	.3026051	.0085037	35.59	0.000	.2859309 .3192794
previous_scores	.0520276	.0035621	14.61	0.000	.045043 .0590122
attendance	.2031513	.0043868	46.31	0.000	.1945495 .211753
sleep_hours	-.0488185	.0349614	-1.40	0.163	-.1173714 .0197344
physical_activity	.1551405	.049265	3.15	0.002	.058541 .25174
_cons	40.94531	.5675232	72.15	0.000	39.8325 42.05811

Регрессия 3. Только юноши

Source	SS	df	MS	Number of obs	=	3,814
Model	31756.3764	5	6351.27528	F(5, 3808)	=	1082.97
Residual	22332.7995	3,808	5.86470576	Prob > F	=	0.0000
				R-squared	=	0.5871
				Adj R-squared	=	0.5866
Total	54089.1759	3,813	14.1854644	Root MSE	=	2.4217

exam_score	Coefficient	Std. err.	t	P> t	[95% conf. interval]
hours_studied	.2811927	.0065533	42.91	0.000	.2683444 .294041
previous_scores	.0443851	.0027141	16.35	0.000	.039064 .0497062
attendance	.1955706	.0034154	57.26	0.000	.1888745 .2022668
sleep_hours	-.0044146	.026594	-0.17	0.868	-.0565545 .0477253
physical_activity	.1558622	.038136	4.09	0.000	.0810932 .2306312
_cons	42.19647	.4326258	97.54	0.000	41.34827 43.04467

Таблица 6. Тест Бокса-Кокса

Test H0:	Restricted log likelihood	LR statistic chi2	Prob > chi2
theta = -1	-26249.023	10099.65	0.000
theta = 0	-21954.673	1510.95	0.000
theta = 1	-21199.383	0.37	0.543

Регрессия 5. С квадратами переменной hours_studied

Source	SS	df	MS	Number of obs	=	6,607
Model	57402.8287	6	9567.13811	F(6, 6600)	=	1482.82
Residual	42583.25	6,600	6.45200758	Prob > F	=	0.0000
Total	99986.0787	6,606	15.1356462	R-squared	=	0.5741
				Adj R-squared	=	0.5737
				Root MSE	=	2.5401

exam_score	Coefficient	Std. err.	t	P> t	[95% conf. interval]
hours_studied	.246944	.0251036	9.84	0.000	.1977328 .2961553
hours_studied2	.0010803	.0006135	1.76	0.078	-.0001224 .0022829
previous_scores	.0475018	.0021727	21.86	0.000	.0432426 .0517611
attendance	.1987083	.0027081	73.37	0.000	.1933995 .2040171
sleep_hours	-.0234534	.0212977	-1.10	0.271	-.0652039 .018297
physical_activity	.1544561	.0303157	5.09	0.000	.0950275 .2138847
_cons	42.08138	.4156432	101.24	0.000	41.26659 42.89618

Регрессия 6. Полулогарифмическая модель

Source	SS	df	MS	Number of obs	=	6,607
Model	12.7669114	5	2.55338228	F(5, 6601)	=	2125.47
Residual	7.92996037	6,601	.001201327	Prob > F	=	0.0000
Total	20.6968718	6,606	.003133041	R-squared	=	0.6169
				Adj R-squared	=	0.6166
				Root MSE	=	.03466

lnexam_score	Coefficient	Std. err.	t	P> t	[95% conf. interval]
hours_studied	.0043311	.0000712	60.82	0.000	.0041915 .0044707
previous_scores	.0007053	.0000296	23.80	0.000	.0006472 .0007634
attendance	.0029645	.000037	80.22	0.000	.0028921 .003037
sleep_hours	-.0002907	.0002906	-1.00	0.317	-.0008604 .000279
physical_activity	.0023948	.0004137	5.79	0.000	.0015839 .0032057
_cons	3.824982	.0047234	809.79	0.000	3.815722 3.834241

Регрессия 7. Линейная в логарифмах модель

Source	SS	df	MS	Number of obs	=	6,561
Model	12.190033	5	2.43800659	F(5, 6555)	=	1915.54
Residual	8.34287995	6,555	.001272751	Prob > F	=	0.0000
Total	20.5329129	6,560	.003130017	R-squared	=	0.5937
				Adj R-squared	=	0.5934
				Root MSE	=	.03568

lnexam_score	Coefficient	Std. err.	t	P> t	[95% conf. interval]
lnhours_studied	.0689752	.0012402	55.62	0.000	.0665441 .0714064
lnprevious_scores	.0517531	.0022369	23.14	0.000	.047368 .0561382
lnattendance	.2338695	.0030104	77.69	0.000	.2279681 .239771
lnsleep_hours	-.0016931	.0020036	-0.85	0.398	-.0056208 .0022347
lnphysical_activity	.005864	.0011381	5.15	0.000	.0036329 .008095
_cons	2.756272	.017437	158.07	0.000	2.72209 2.790455