

ПОЛИТЕХ
Санкт-Петербургский
политехнический университет
Петра Великого



Институт дополнительного
образования



ВЫСШАЯ
ИНЖЕНЕРНАЯ
ШКОЛА

Выпускная квалификационная работа

Прогнозирование ВВП с помощью методов машинного обучения

Выполнил: Григорьев Ярослав Владимирович

Руководитель: Заграновская Анна Васильевна

Анализ данных на языке Python

Проблема

Официальные статистические данные о годовой динамике ВВП публикуются с существенным запаздыванием, однако для аналитических исследований и прогнозирования результатов выполненных действий, требуется своевременное получение качественных данных о показателях ВВП.

Москва. 18 февраля. INTERFAX.RU - ВВП РФ в 2021 году вырос на 4,7%. Такую первую оценку опубликовал в пятницу Росстат. Ранее Минэкономразвития [оценило](#) рост ВВП РФ в 2021 году на уровне 4,6% по сравнению с 2020 годом.

интерфакс

ЭКОНОМИКА 19:19, 18 февраля 2022



Цель

Найти наиболее точную модель прогнозирования ВВП методами машинного обучения.

Задачи

1. Построить прогноз ВВП на основе производственной функции Кобба— Дугласа, где Капитал и Труд спрогнозировать на основе моделей авторегрессии.
2. Построить прогноз ВВП на основе производственной функции Кобба— Дугласа, где Капитал и Труд спрогнозировать на основе моделей модели Хольта-Уинтерса и Тейла-Вейджа.
3. Построить прогноз ВВП на основе адаптивных тренд-сезонных моделей(модели Хольта-Уинтерса и Тейла-Вейджа).
4. Построить прогноз ВВП на основе модели регрессии на факторные признаки Капитал и Труд и их лагированные значения.

Источники данных

Для анализа и прогнозирования были взяты данные Внутреннего Валового Продукта (ВВП) Российской Федерации за период с 1990 по 2020 гг.

Также данные о показателях Капитала и Труда как факторы, непосредственно влияющие на объем производства за тот же период.

Данные за период с 1990 по 2020 гг. для обучения получены с сайта Всемирного банка:

1. URL: <http://data.worldbank.org/indicator/NY.GDP.MKTP.CD?view=chart>
2. URL: <http://data.worldbank.org/indicator/NE.GDI.TOTL.CD?view=chart>
3. URL: <http://data.worldbank.org/indicator/SL.TLF.TOTL.IN?view=chart>

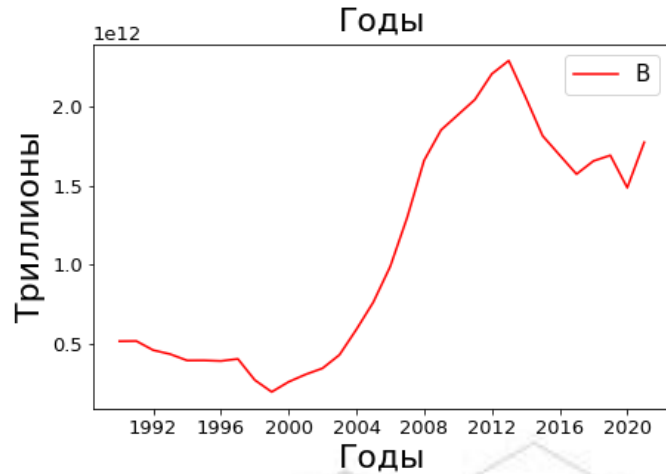
Данные за период 2019-2022гг. для проверки модели взяты с сайта Международного валютного фонда:

1. URL: <https://www.imf.org/en/Countries/RUS#countrydata>

Предварительная обработка данных

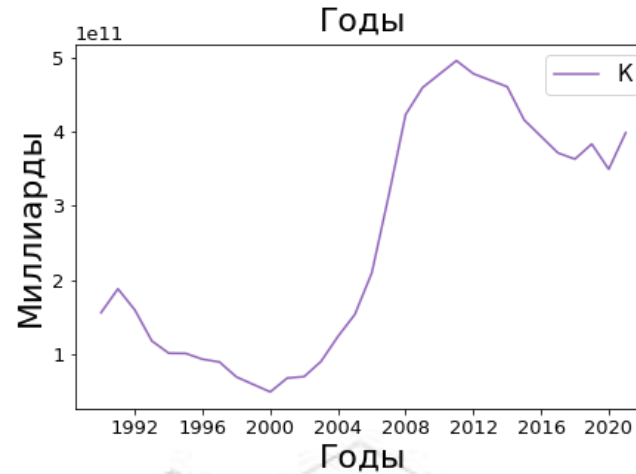
ВВП

Исходная и сглаженная



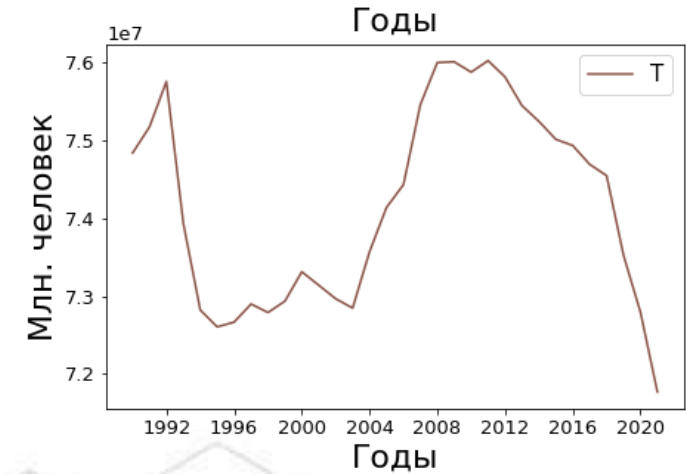
Капитал

Исходная и сглаженная



Труд

Исходная и сглаженная



Построение прогноза на основе производственной функции Кобба— Дугласа

Прогнозирование Капитала, Труда и на их основе - ВВП.



Производственная функция Кобба-Дугласа

$$Y = AK^{\alpha}N^{\beta}$$

где A — технологический коэффициент,
 α — коэффициент эластичности по труду,
 β — коэффициент эластичности по капиталу.

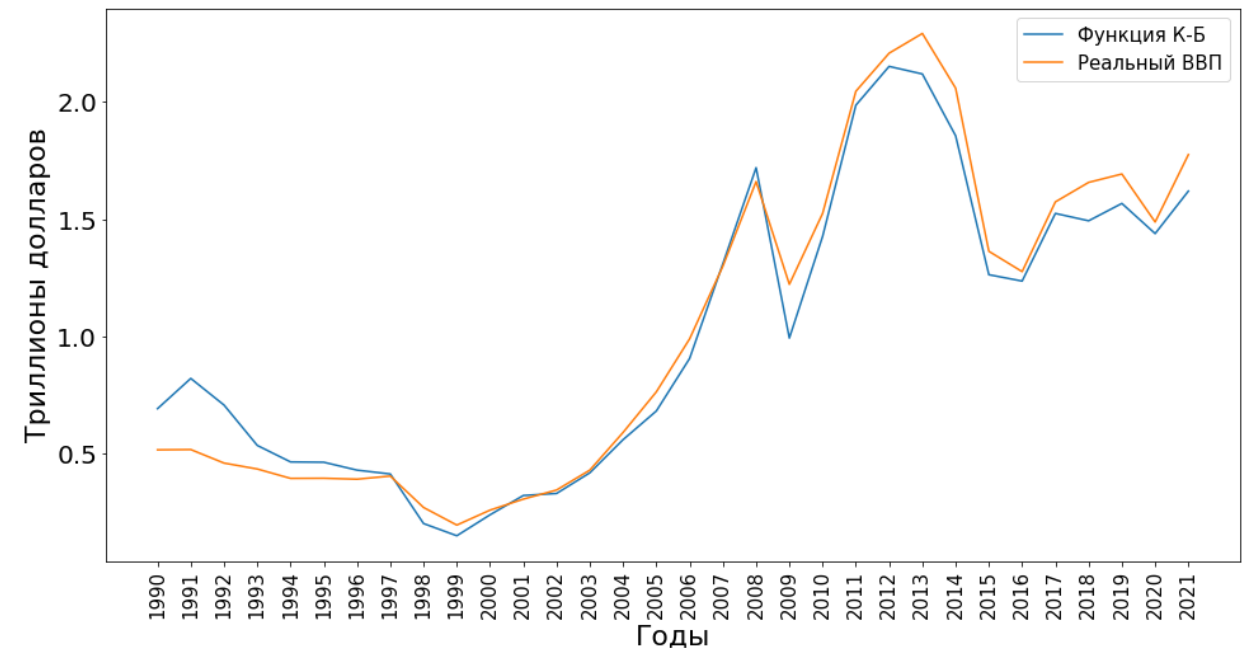
$R^2 = 0.96$

$MAE = 93444637954.33\$$

$MSE = 1.479e+22$

$RMSE = 119157843671.47\$$

Расчетные значения по модели Кобба-Дугласа



Прогнозирование значений Капитала и Труда на основе моделей авторегрессии



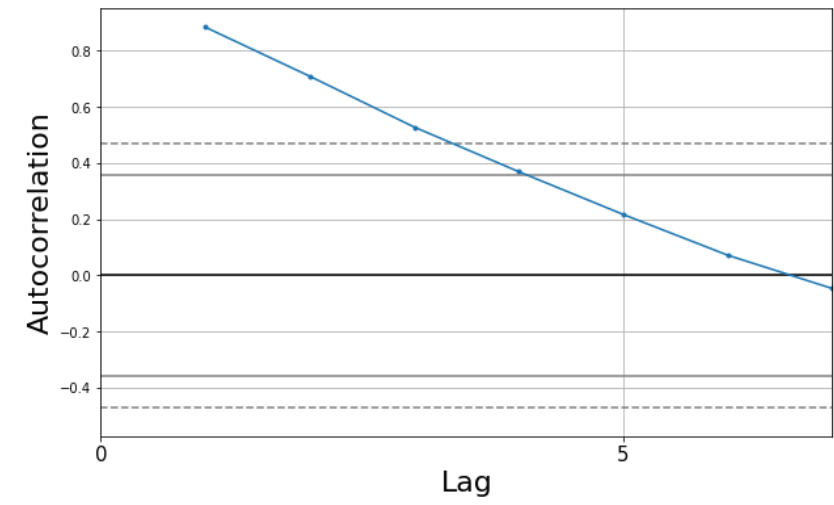
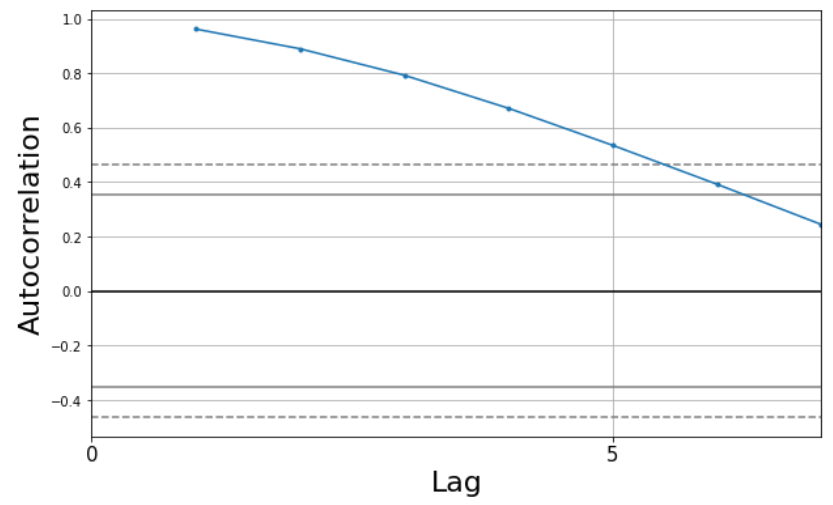
ВВП = f(Капитал, Труд)

Капитал

Труд

Авторегрессионная модель порядка 5
 $K_t = f(K_{t-1}, K_{t-2}, \dots, K_{t-5})$

Авторегрессионная модель порядка 3
 $L_t = f(L_{t-1}, L_{t-2}, L_{t-3})$



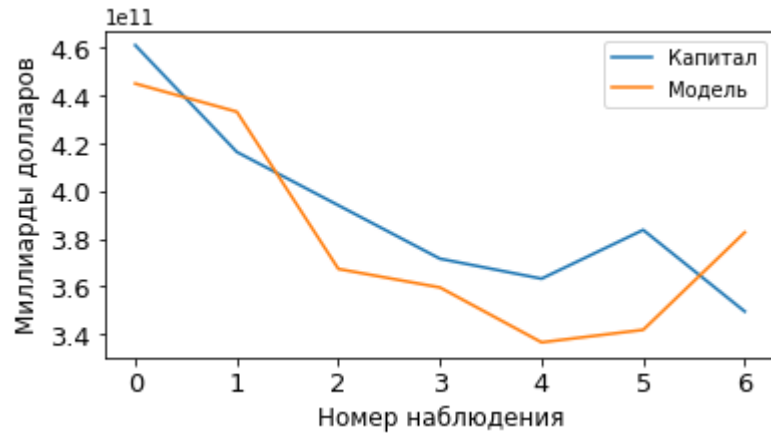
Используемые модели:

- sklearn.linear_model LinearRegression,
- sklearn.linear_model Ridge,
- xgboost.XGBRegressor

model = Ridge()

Капитал

Проверка модели на тестовых данных



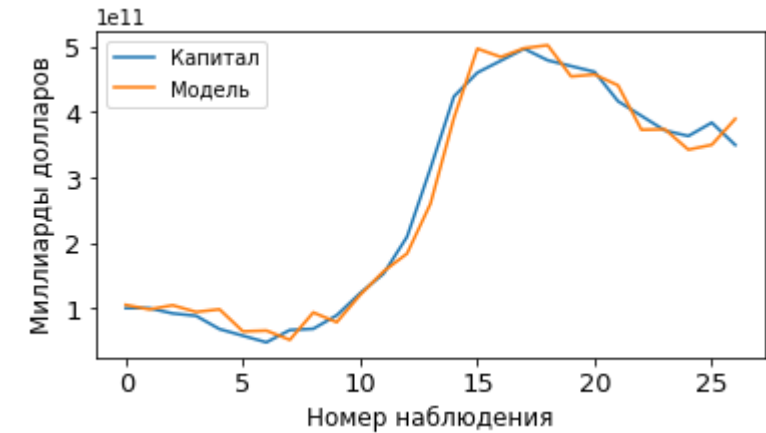
$$R^2 = 0.41$$

$$MAE = 24758271965.14\$$$

$$MSE = 7.086e+20$$

$$RMSE = 26619730531.32\$$$

Расчетные значения по модели



$$R^2 = 0.98$$

$$MAE = 17626795666.45\$$$

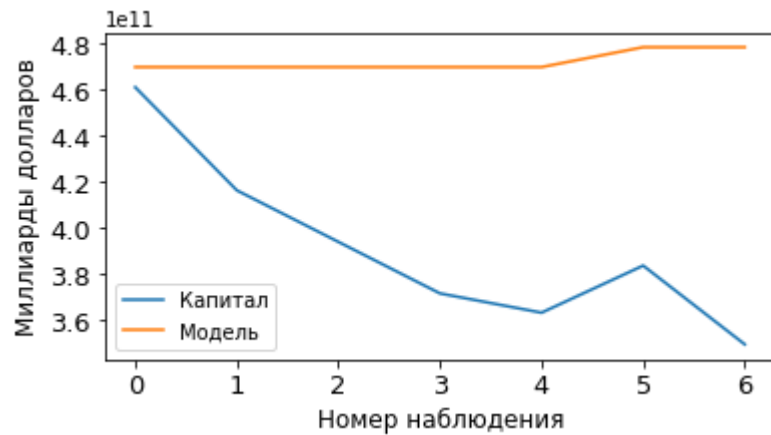
$$MSE = 4.988e+20$$

$$RMSE = 22335386994.66\$$$

model = xgboost.XGBRegressor()

Капитал

Проверка модели на тестовых данных



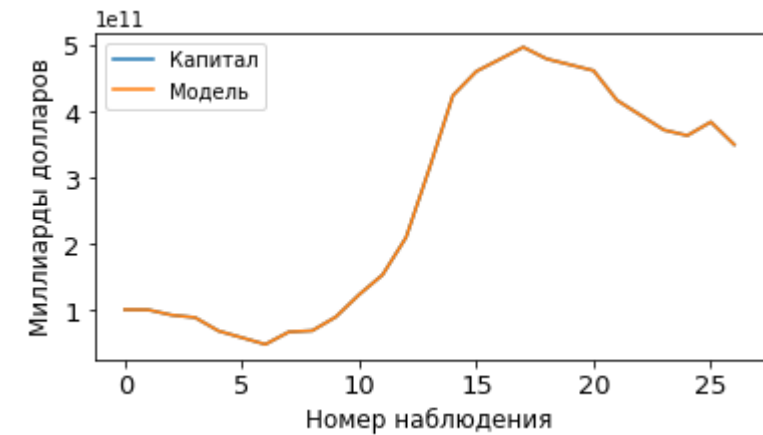
$$R^2 = -5.52$$

$$MAE = 80950179576.96\$$$

$$MSE = 7.901e+21$$

$$RMSE = 88892729849.50\$$$

Расчетные значения по модели



$$R^2 = 0.99$$

$$MAE = 173922.50\$$$

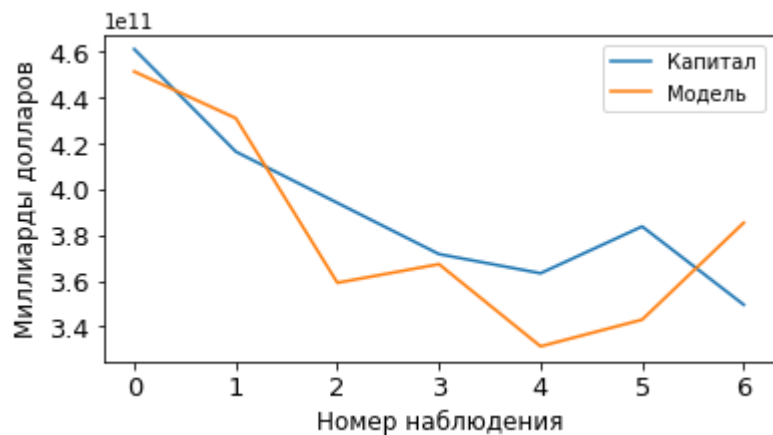
$$MSE = 50121745031.90$$

$$RMSE = 223878.87\$$$

model = LinearRegression()

Капитал

Проверка модели на тестовых данных



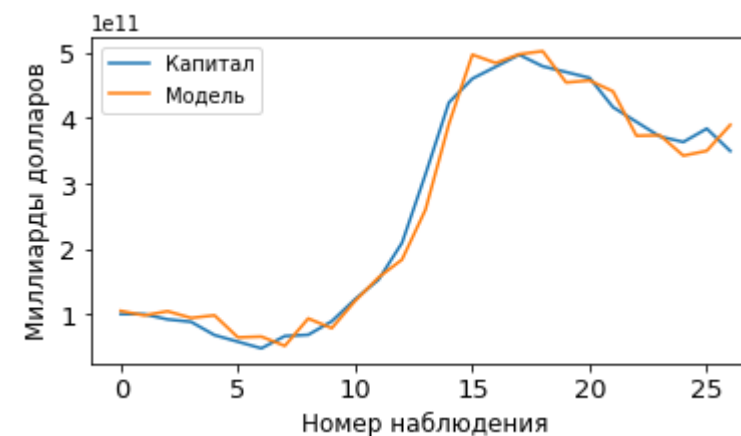
$$R^2 = 0.35$$

$$MAE = 24575252715.49\$$$

$$MSE = 7.855e+20$$

$$RMSE = 28027664165.73\$$$

Расчетные значения по модели

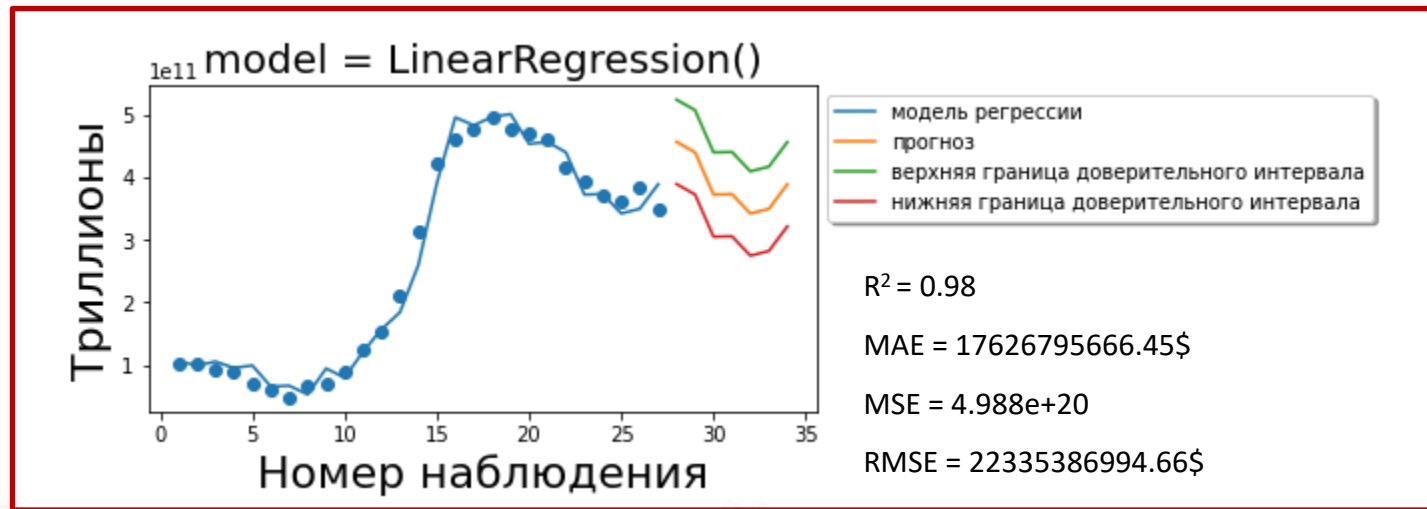
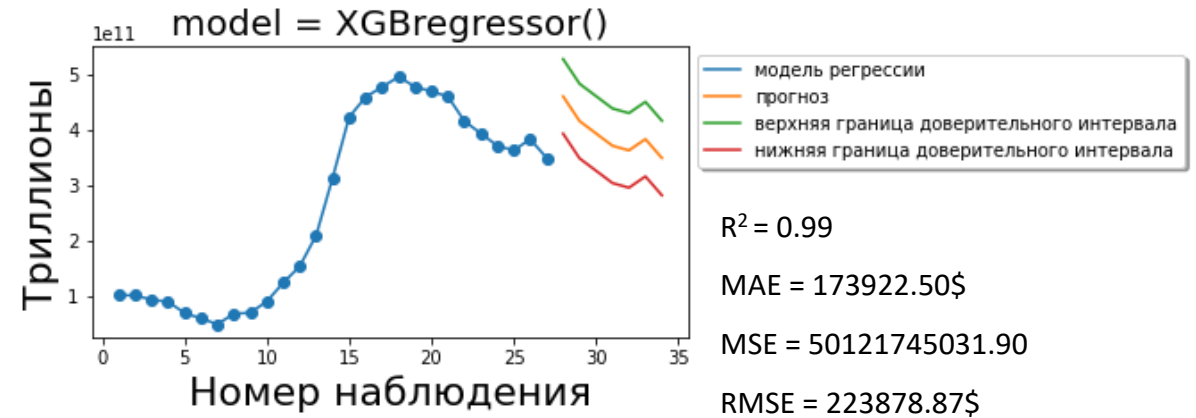
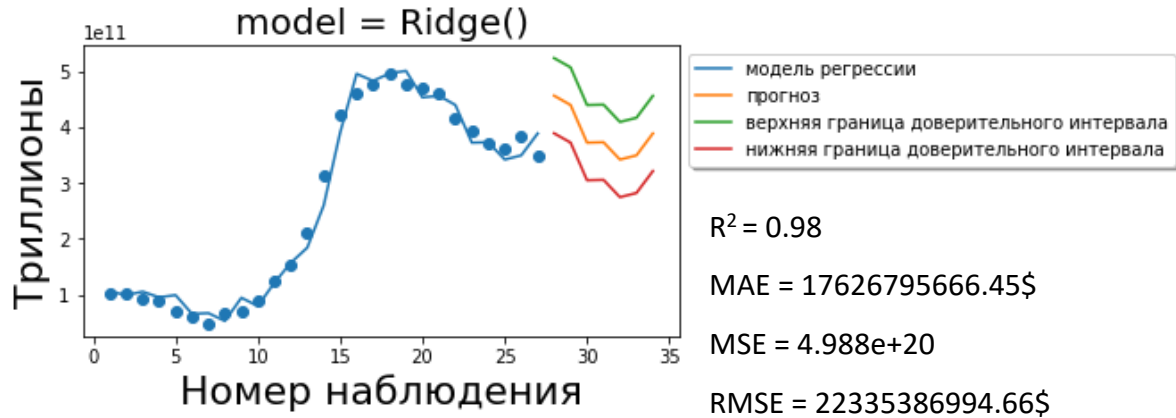


$$R^2 = 0.98$$

$$MAE = 17626795666.45\$$$

$$MSE = 4.988e+20$$

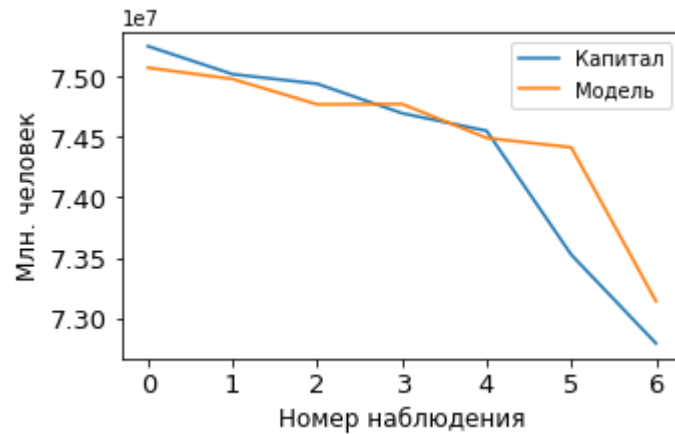
$$RMSE = 22335386994.65\$$$



model = Ridge()

Труд

Проверка модели на тестовых данных



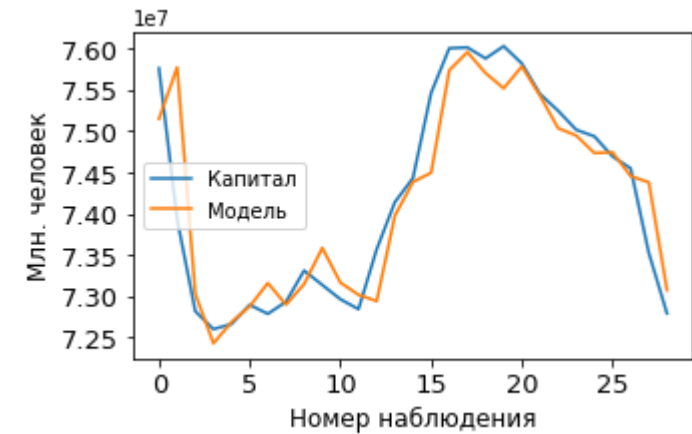
$$R^2 = 0.79$$

MAE = 251107.39 человек

MSE = 139800565563.68

RMSE = 373899.14 человек

Расчетные значения по модели



$$R^2 = 0.84$$

MAE = 306782.48 человек

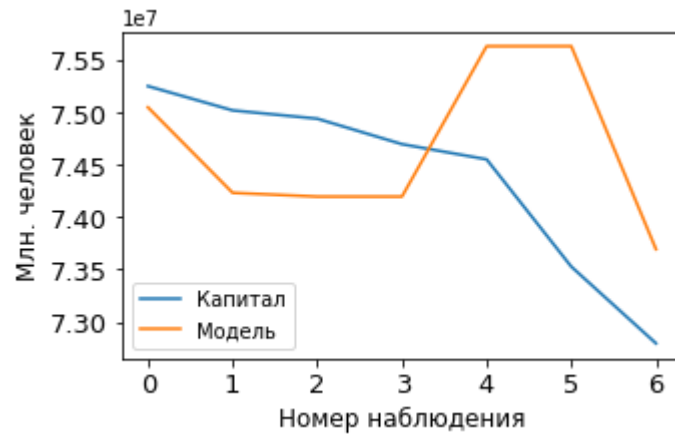
MSE = 237766746566.90

RMSE = 487613.32 человек

model = xgboost.XGBRegressor()

Труд

Проверка модели на тестовых данных



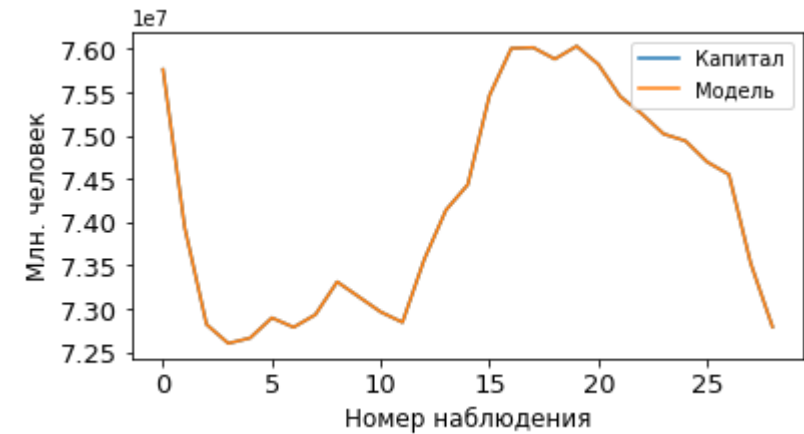
$$R^2 = -0.62$$

MAE = 902184.29 человек

MSE = 1123108666635.42

RMSE = 1059768.21 человек

Расчетные значения по модели



$$R^2 = 0.99$$

MAE = 65.07 человек

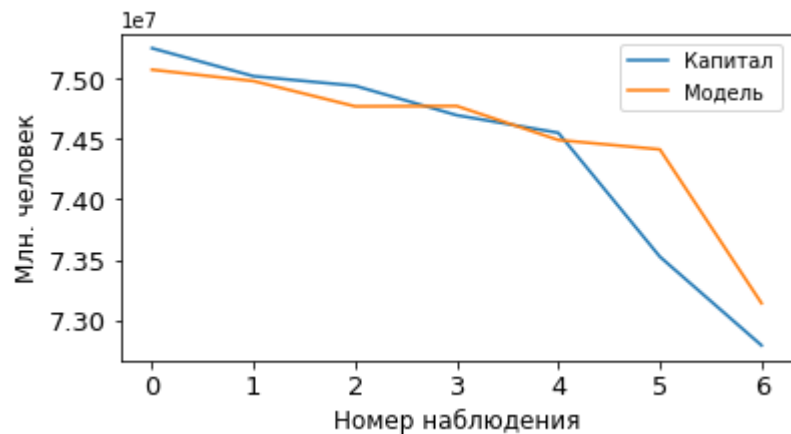
MSE = 6153.06

RMSE = 78.44 человек

model = LinearRegression()

Труд

Проверка модели на тестовых данных



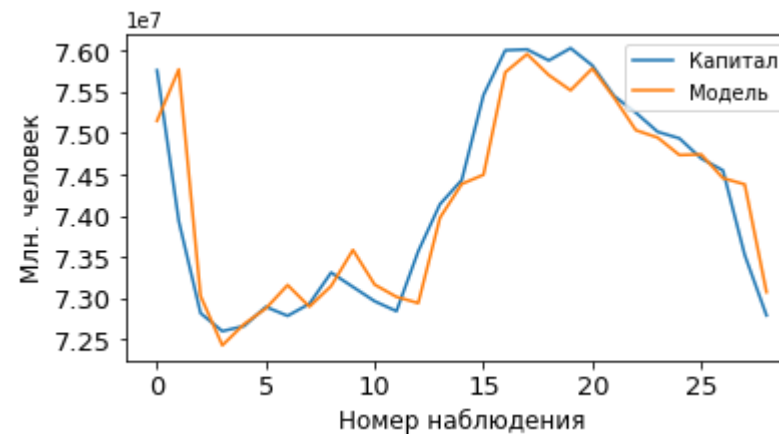
$$R^2 = 0.79$$

MAE = 251107.39 человек

MSE = 139800565563.66

RMSE = 373899.14 человек

Расчетные значения по модели



$$R^2 = 0.84$$

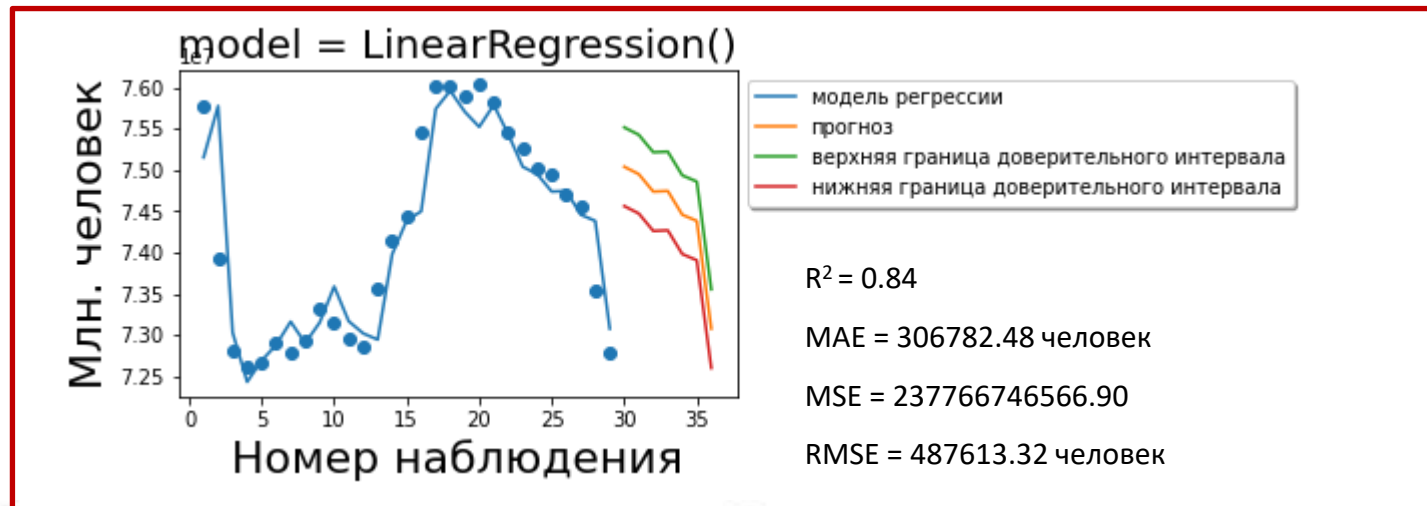
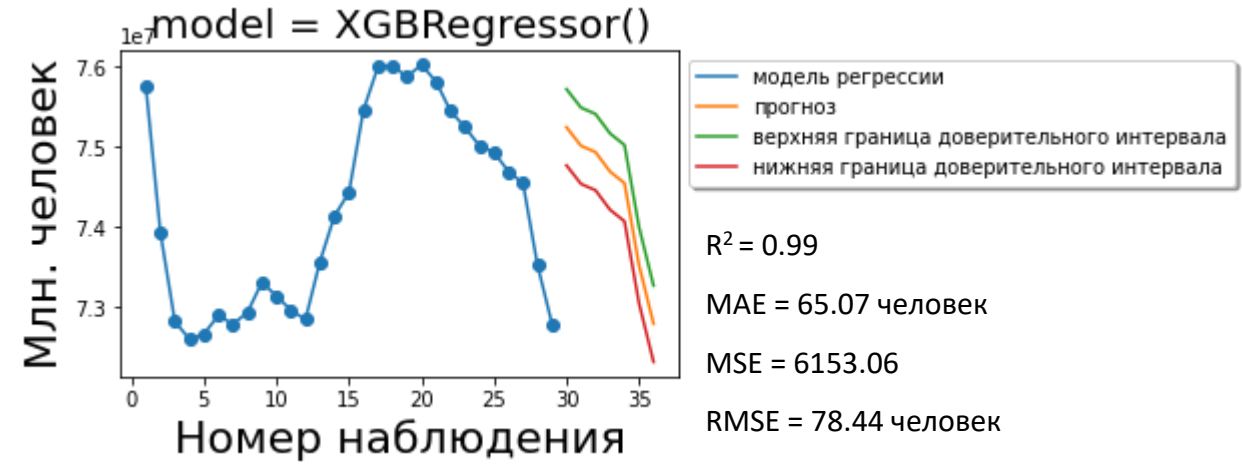
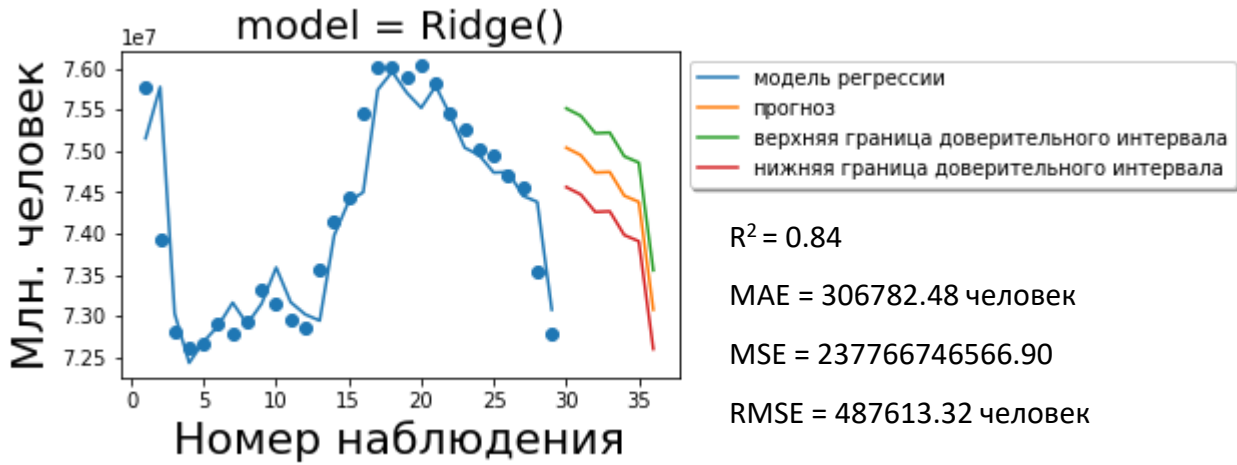
MAE = 306782.48 человек

MSE = 237766746566.90

RMSE = 487613.32 человек

Труд

Прогнозы



Прогноз по производственной функции

```
1 def cobb_douglas_func_pred(dataframe, x, y):
2     import pandas as pd
3     import numpy as np
4     df = pd.DataFrame({
5         'Y': np.nan,
6         'K': dataframe[x],
7         'L': dataframe[y]
8     })
9     Y = 9.375 * np.power(df.K, 0.902) * np.power(df.L, 0.098)
10    return Y
```

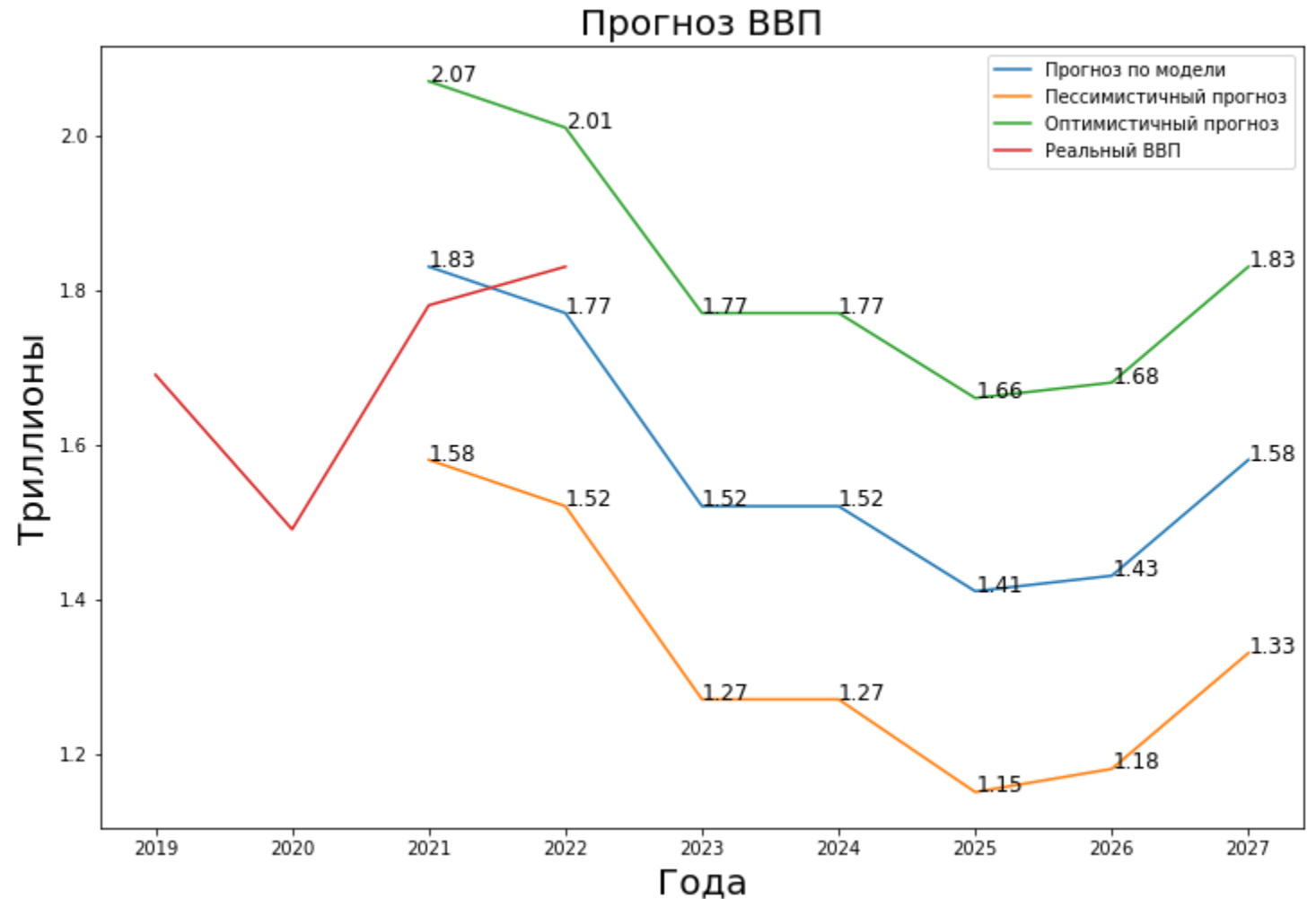
Прогноз на основании производственной функции Кобба-Дугласа.

$R^2 = 0.95$

MAE = 128095480743.00\$

MSE = 2.618e+22

RMSE = 161802925312.92\$

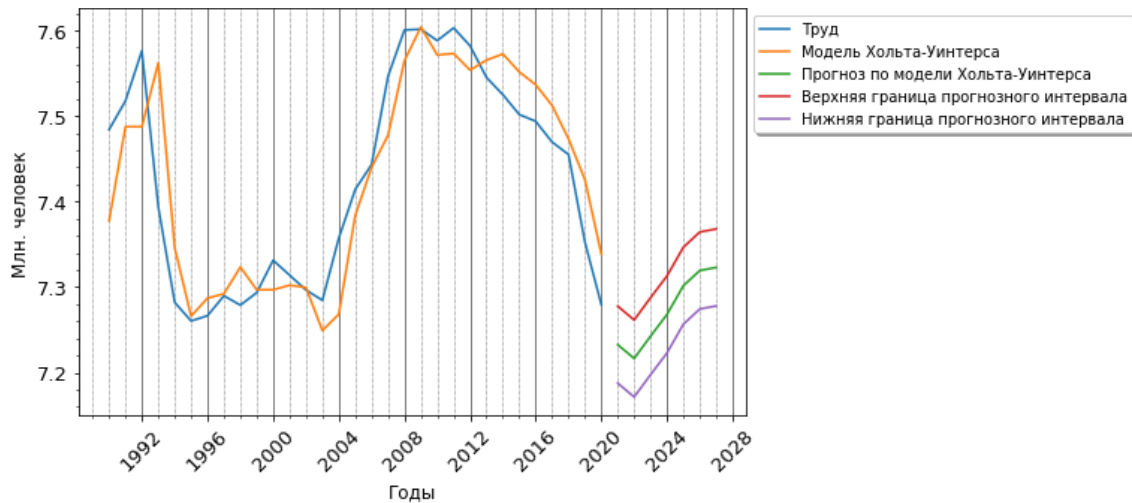


Прогнозирование значений Капитала и Труда на основе моделей Хольта-Уинтерса и Тейла- Вейджа



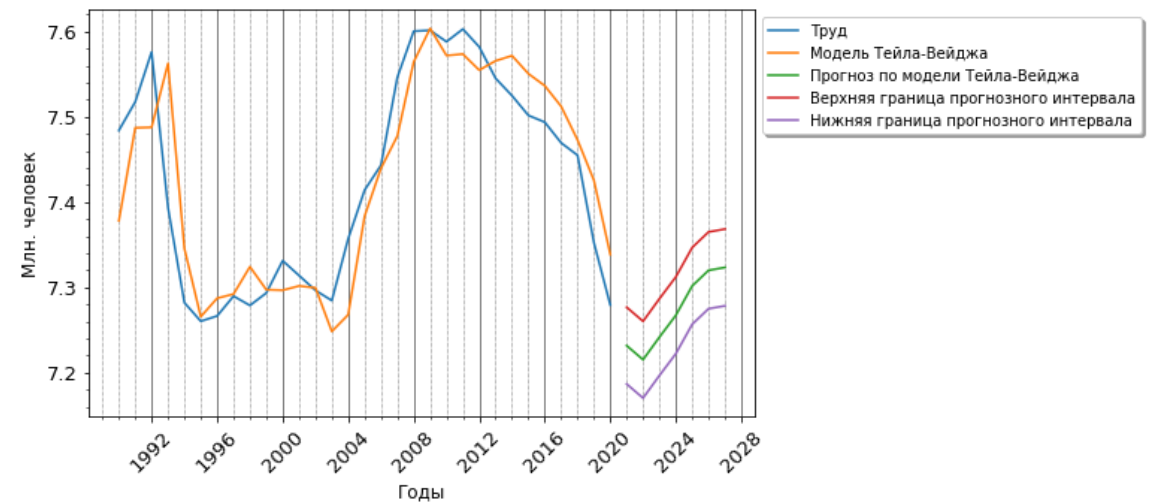
Прогнозирование Труда на основе моделей Хольта-Уинтерса и Тейла-Вейджа

Модель Хольта-Уинтерса



MAE = 412097.79 человек
MSE = 298467740440.6458
RMSE = 546322.01 человек
 $R^2 = 0.79$

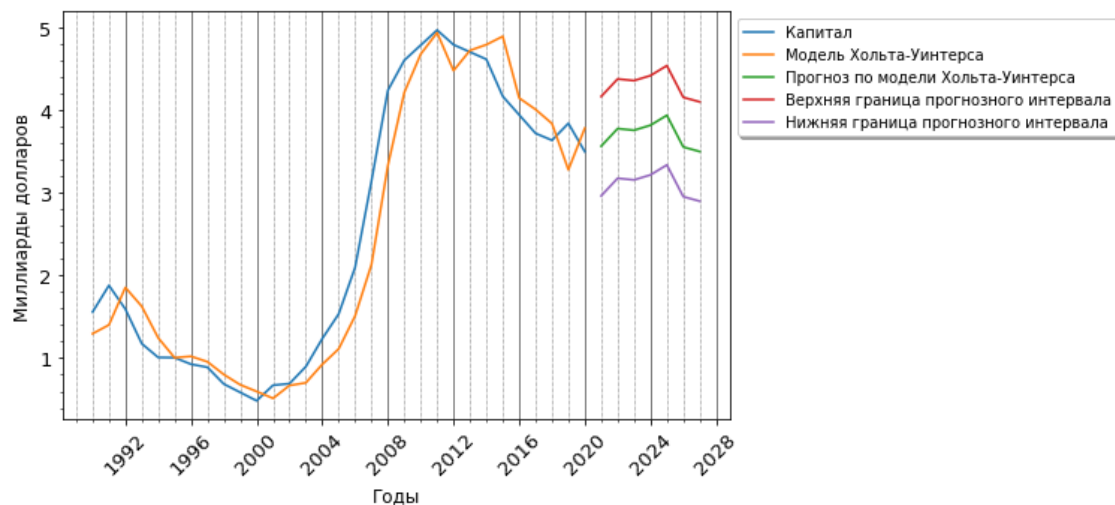
Модель Тейла-Вейджа



MAE = 411858.43 человек
MSE = 298290037163.16
RMSE = 546159.35 человек
 $R^2 = 0.79$

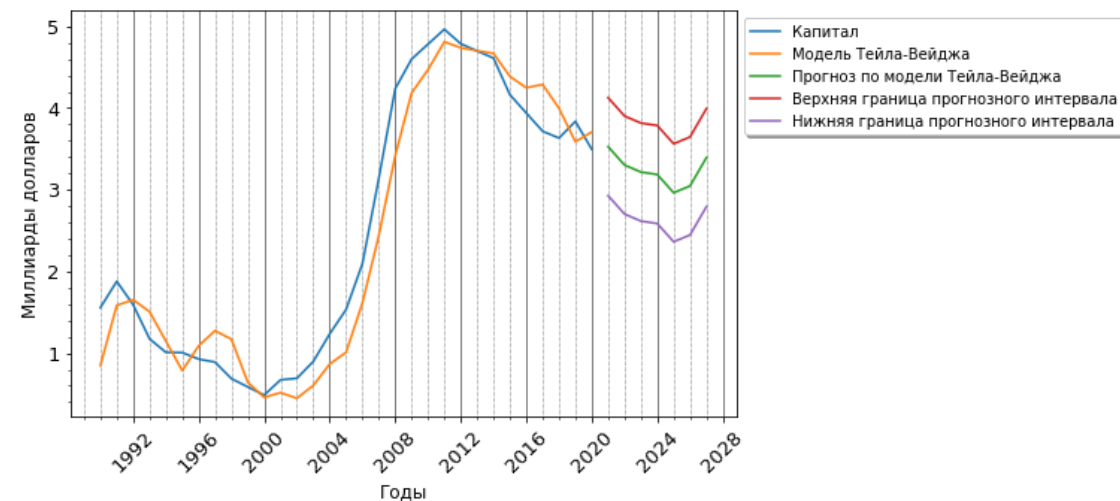
Прогнозирование Капитала на основе моделей Хольта-Уинтерса и Тейла-Вейджа

Модель Хольта-Уинтерса



MAE = 29289485998.22\$
MSE = 1.475e+21
RMSE = 38415085733.49\$
 $R^2 = 0.94$

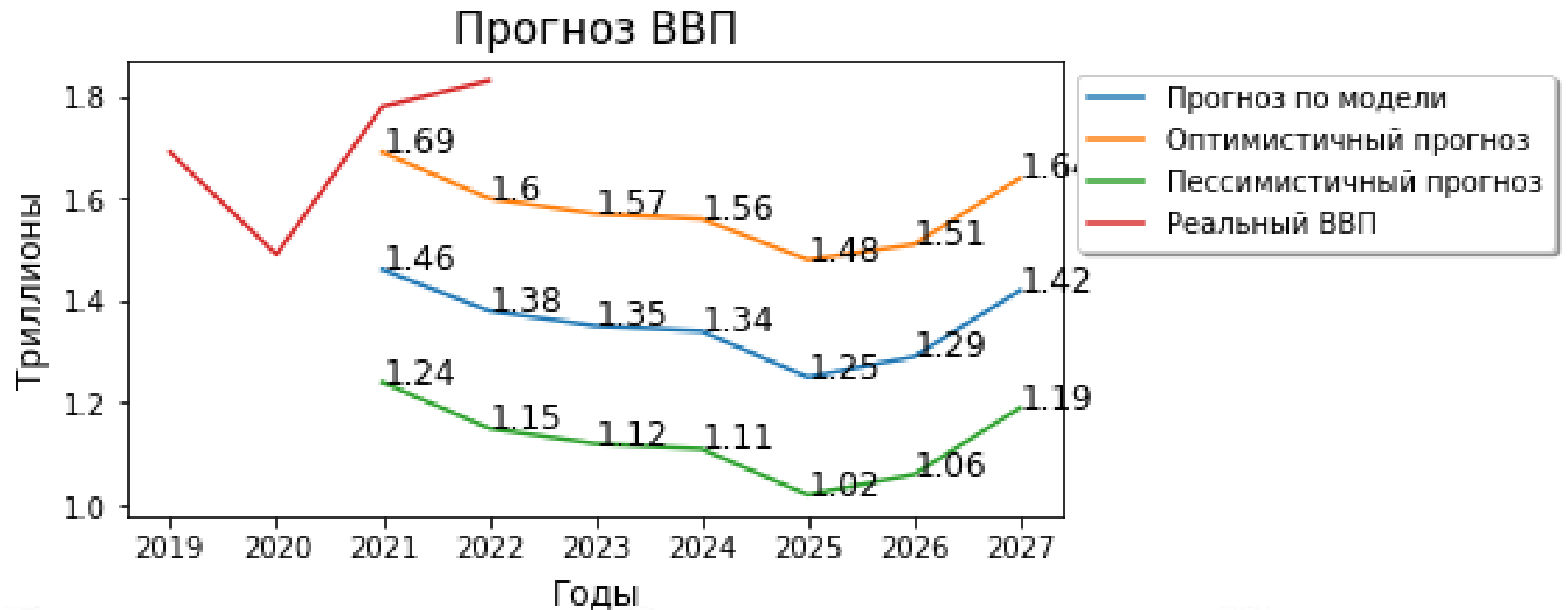
Модель Тейла-Вейджа



MAE = 30243579020.31\$
MSE = 1.344e+21
RMSE = 36672592787.53\$
 $R^2 = 0.94$

Производственная функция Кобба-Дугласа

Модель Тейла-Вейджа



Построение прогноза ВВП на основе адаптивных тренд- сезонных моделей

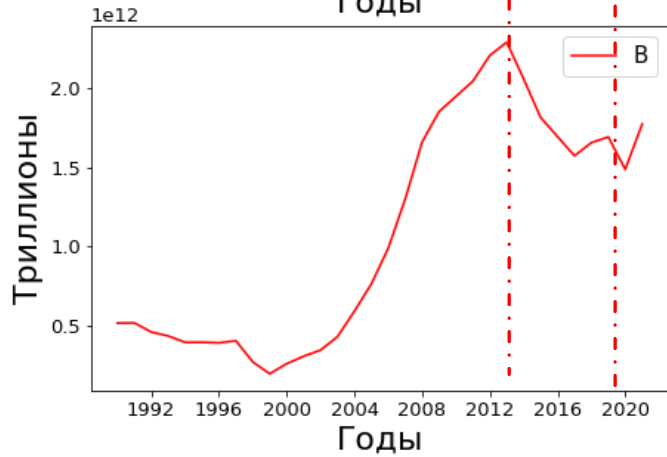
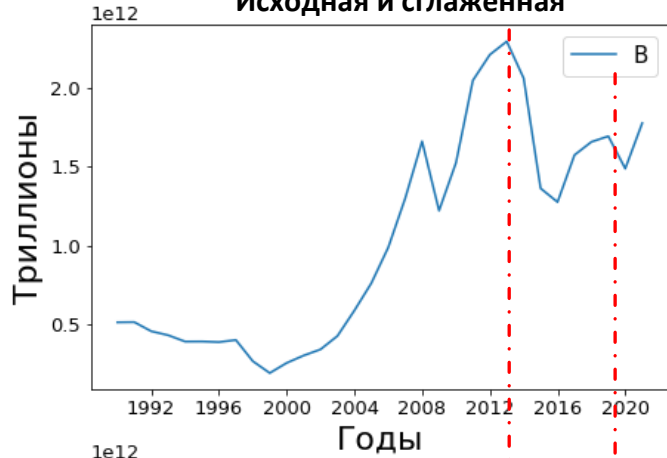
Модели Хольта-Уинтерса и Тейла-Вейджа



Визуальное определение лага цикличности

ВВП

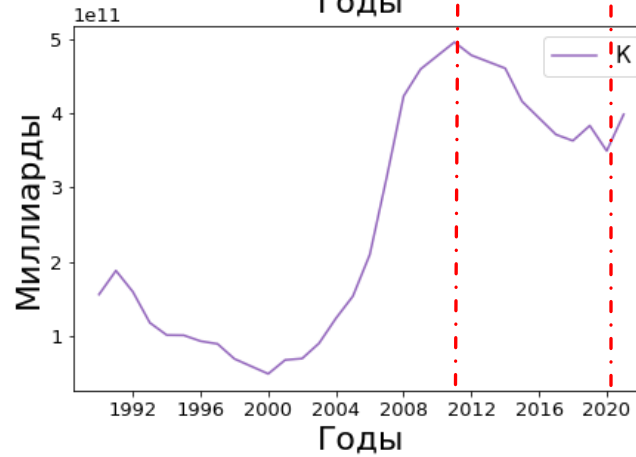
Исходная и сглаженная



$s \approx 7-8$ лет

Капитал

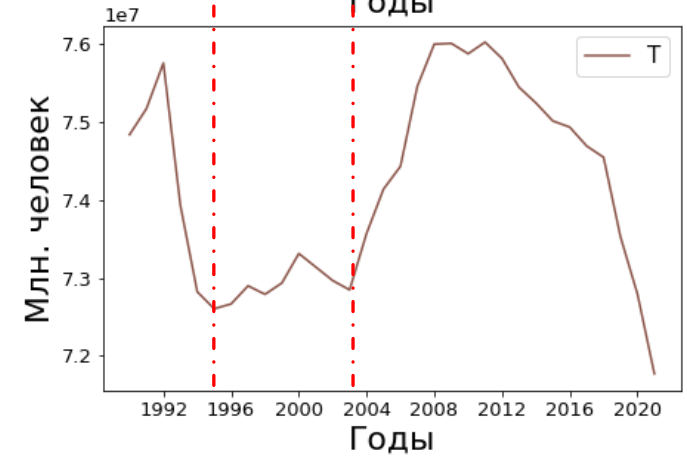
Исходная и сглаженная



$s \approx 7-8$ лет

Труд

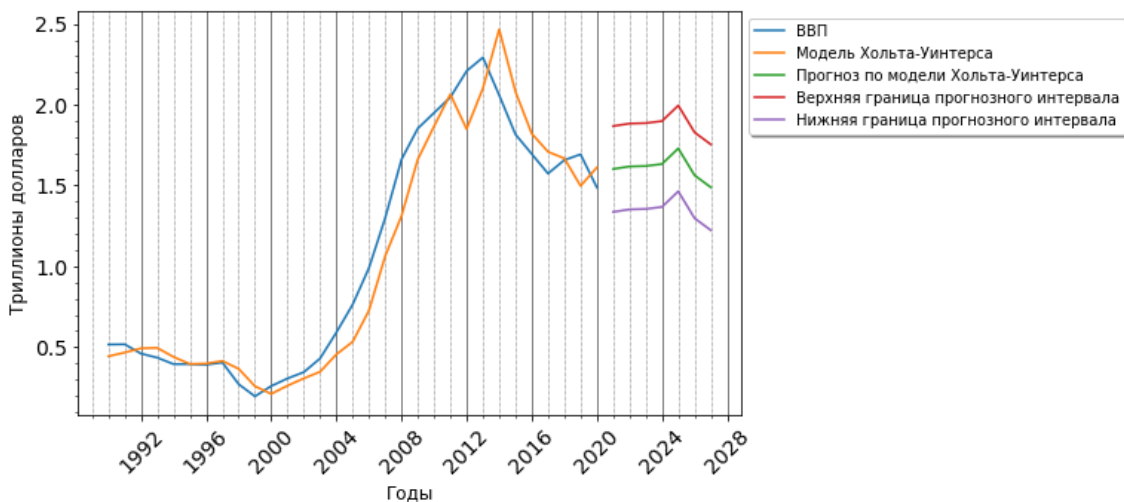
Исходная и сглаженная



$s \approx 8-9$ лет

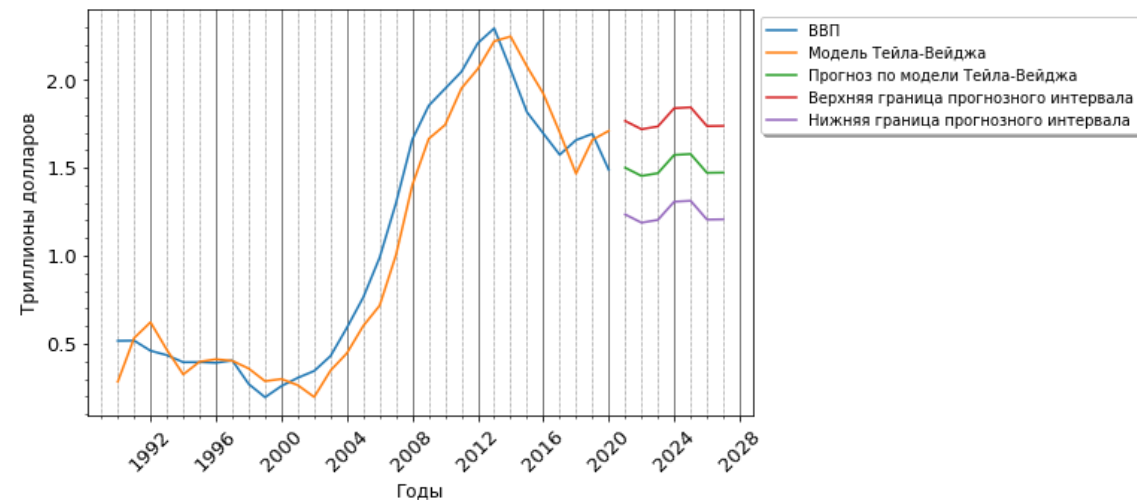
Модели Хольта-Уинтерса и Тейла-Вейджа для ВВП

Модель Хольта-Уинтерса



MAE = 128067946399.90\$
MSE = 2.874e+22
RMSE = 169548321755.58\$
 $R^2 = 0.94$

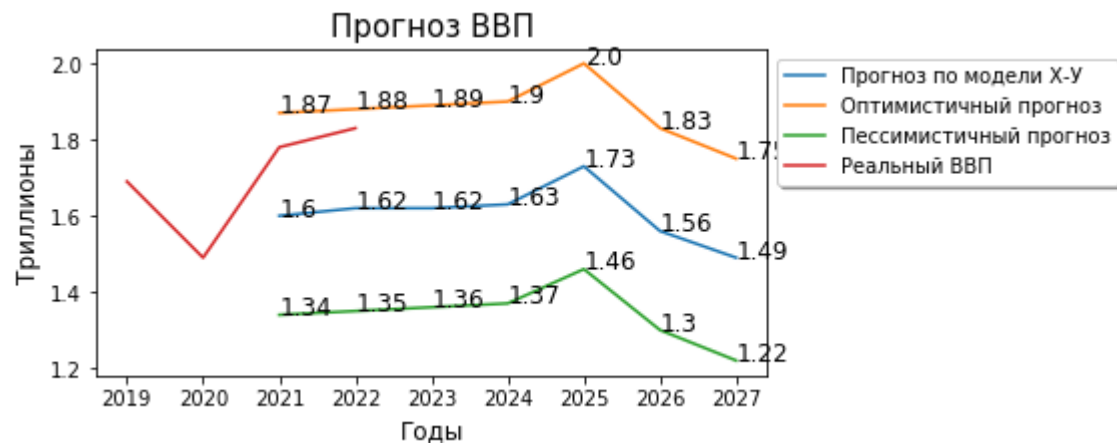
Модель Тейла-Вейджа



MAE = 132767408003.96\$
MSE = 2.519e+22
RMSE = 158727840509.01\$
 $R^2 = 0.95$

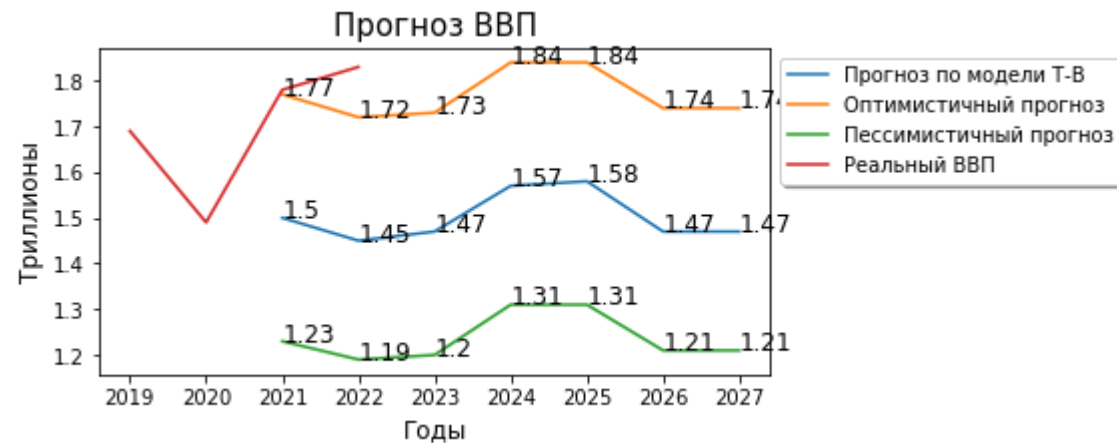
Прогнозы ВВП по моделям Хольта-Уинтерса и Тейла-Вейджа

Модель Хольта-Уинтерса



MAE = 128067946399.90\$
MSE = 2.874e+22
RMSE = 169548321755.58\$
 $R^2 = 0.94$

Модель Тейла-Вейджа



MAE = 132767408003.96\$
MSE = 2.519e+22
RMSE = 158727840509.01\$
 $R^2 = 0.95$

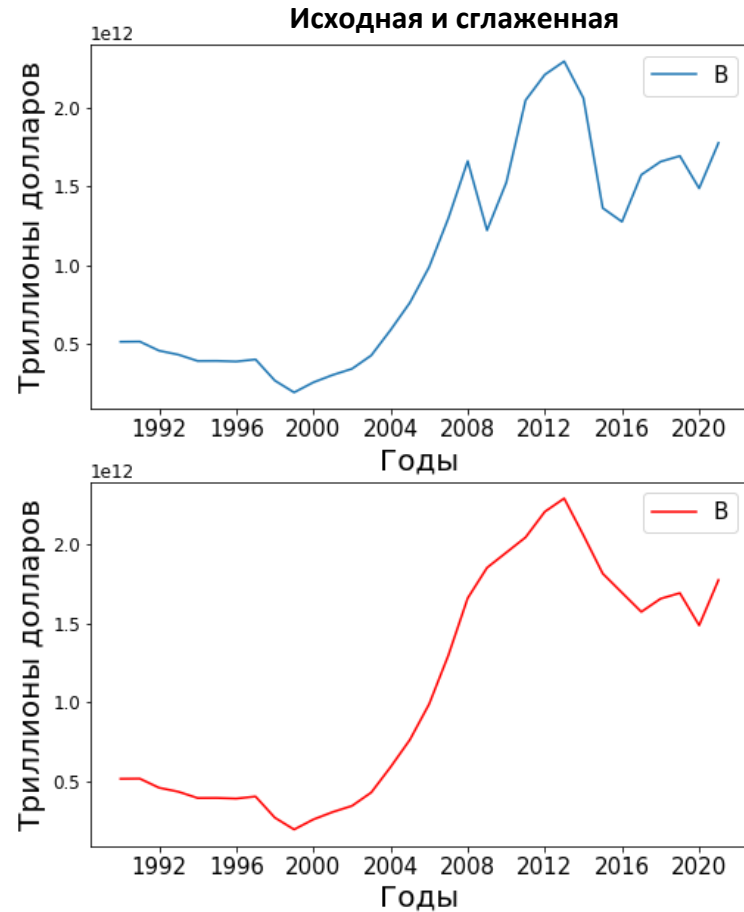
Построение прогноза ВВП на основе модели регрессии с рядом факторов, и их лагированными значениями.

В качестве факторных признаков рассматриваются капитал и труд с лагами 1-6

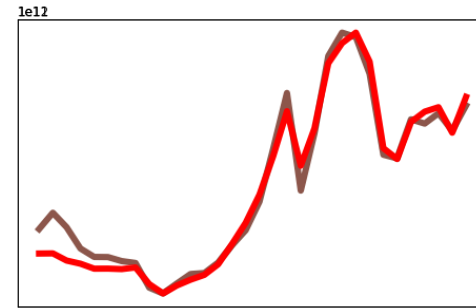


Схожесть тенденции у Капитала и ВВП

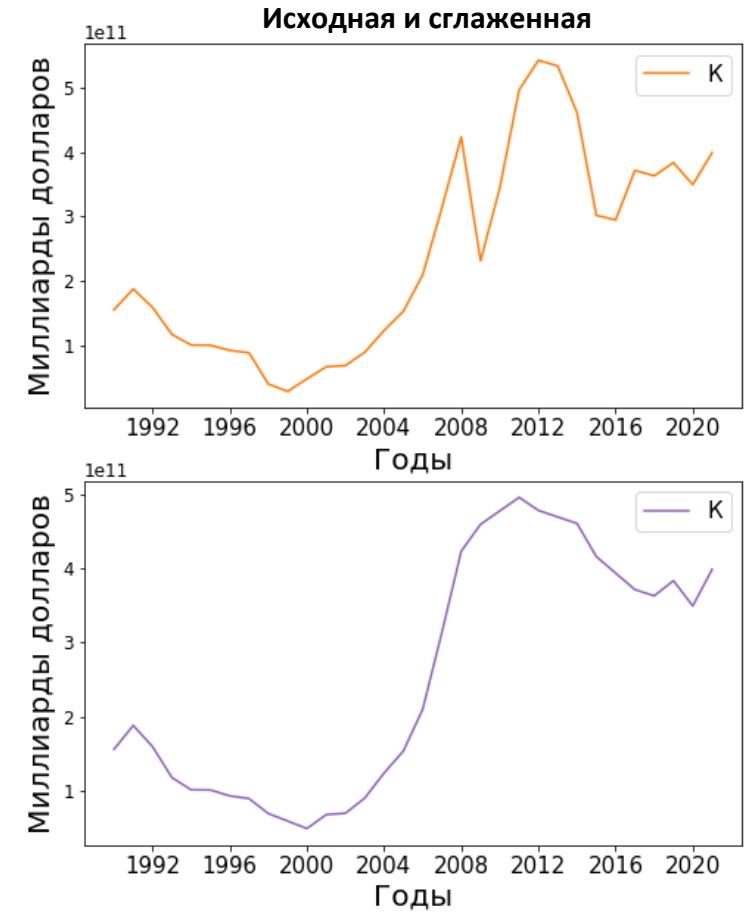
ВВП



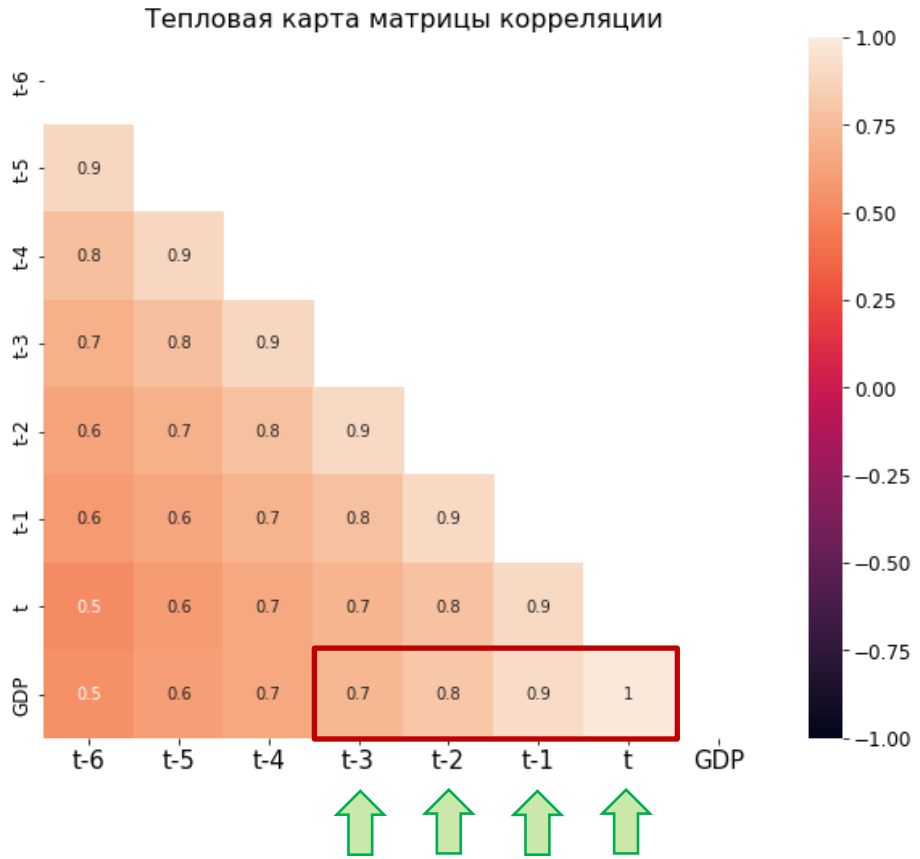
— ВВП
— Капитал



Капитал



Коррелограмма для ВВП, Капитала и его лагированных значений



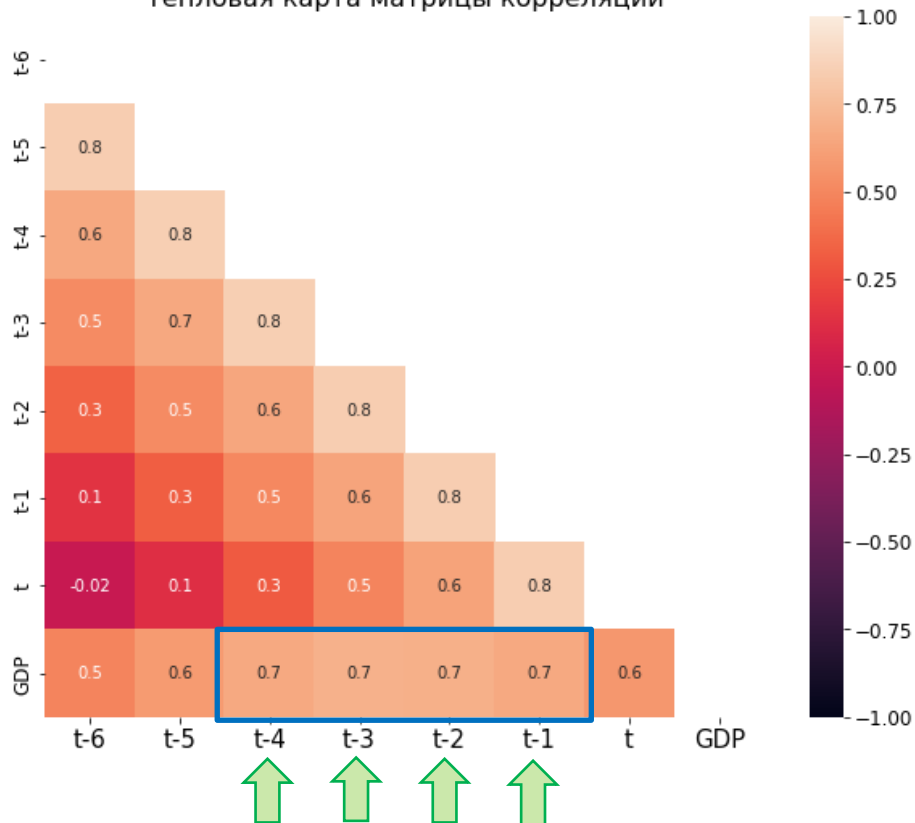
Матрица корреляции

	t-6	t-5	t-4	t-3	t-2	t-1	t	GDP
t-6	1.00000	0.894382	0.766504	0.695440	0.637525	0.589106	0.520300	0.540458
t-5	0.894382	1.00000	0.893197	0.768814	0.698830	0.641894	0.594694	0.605488
t-4	0.766504	0.893197	1.00000	0.896261	0.776448	0.704883	0.653477	0.658986
t-3	0.695440	0.768814	0.896261	1.00000	0.899301	0.780862	0.713612	0.726704
t-2	0.637525	0.698830	0.776448	0.899301	1.00000	0.900647	0.788057	0.811492
t-1	0.589106	0.641894	0.704883	0.780862	0.900647	1.00000	0.902108	0.916664
t	0.520300	0.594694	0.653477	0.713612	0.788057	0.902108	1.00000	0.988117
GDP	0.540458	0.605488	0.658986	0.726704	0.811492	0.916664	0.988117	1.00000

Лагированные значения Капитала показали высокую степень значимости с лагом до t-3

Коррелограмма для ВВП, Труда и его лагированных значений

Тепловая карта матрицы корреляции



Матрица корреляции

	t-6	t-5	t-4	t-3	t-2	t-1	t	GDP
t-6	1.000000	0.843664	0.649543	0.523352	0.340708	0.145325	-0.021740	0.491334
t-5	0.843664	1.000000	0.844769	0.651708	0.517673	0.327320	0.120904	0.592657
t-4	0.649543	0.844769	1.000000	0.845473	0.646957	0.506314	0.305811	0.657825
t-3	0.523352	0.651708	0.845473	1.000000	0.841168	0.636334	0.484673	0.697079
t-2	0.340708	0.517673	0.646957	0.841168	1.000000	0.840002	0.631450	0.692506
t-1	0.145325	0.327320	0.506314	0.636334	0.840002	1.000000	0.839996	0.661558
t	-0.021740	0.120904	0.305811	0.484673	0.631450	0.839996	1.000000	0.573495
GDP	0.491334	0.592657	0.657825	0.697079	0.692506	0.661558	0.573495	1.000000

Лагированные значения Труда показали достаточную степень значимости с лагом t-1, t-2 и t-3

Модель регрессии ВВП на Труд и его лагированные значения

Полные данные по Труд

OLS Regression Results

Dep. Variable:	GDP	R-squared:	0.846
Model:	OLS	Adj. R-squared:	0.786
Method:	Least Squares	F-statistic:	14.16
Date:	Wed, 06 Jul 2022	Prob (F-statistic):	3.75e-06
Time:	11:24:32	Log-Likelihood:	-720.63
No. Observations:	26	AIC:	1457.
Df Residuals:	18	BIC:	1467.
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	-2.765e+13	3.39e+12	-8.164	0.000	-3.48e+13	-2.05e+13
t-6	5.346e+04	6.52e+04	0.820	0.423	-8.35e+04	1.9e+05
t-5	3.408e+04	9.4e+04	0.363	0.721	-1.63e+05	2.31e+05
t-4	7444.4013	9.86e+04	0.076	0.941	-2e+05	2.15e+05
t-3	8.966e+04	1.05e+05	0.855	0.404	-1.31e+05	3.1e+05
t-2	6.257e+04	1.01e+05	0.620	0.543	-1.5e+05	2.75e+05
t-1	8.926e+04	9.84e+04	0.907	0.376	-1.17e+05	2.96e+05
t	5.387e+04	6.99e+04	0.771	0.451	-9.3e+04	2.01e+05

Omnibus:	0.432	Durbin-Watson:	0.963
Prob(Omnibus):	0.806	Jarque-Bera (JB):	0.010
Skew:	0.022	Prob(JB):	0.995
Kurtosis:	3.088	Cond. No.	1.06e+10

GDP – ВВП

t-6 – Значения Труда с лагом t-6

t-5 – Значения Труда с лагом t-5

t-4 – Значения Труда с лагом t-4

t-3 – Значения Труда с лагом t-3

t-2 – Значения Труда с лагом t-2

t-1 – Значения Труда с лагом t-1

t – Значения Труда

P>|t| – Коэффициент значимости

Adj. R-squared - Скорректированный коэффициент детерминации

Модель регрессии ВВП на Труд и его лагированные значения

Данные после удаления наименее значимых признаков

OLS Regression Results

Dep. Variable:	GDP	R-squared:	0.809			
Model:	OLS	Adj. R-squared:	0.792			
Method:	Least Squares	F-statistic:	48.62			
Date:	Wed, 06 Jul 2022	Prob (F-statistic):	5.48e-09			
Time:	11:26:02	Log-Likelihood:	-723.47			
No. Observations:	26	AIC:	1453.			
Df Residuals:	23	BIC:	1457.			
Df Model:	2					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
const	-2.366e+13	2.53e+12	-9.363	0.000	-2.89e+13	-1.84e+13
t-3	1.966e+05	4.35e+04	4.523	0.000	1.07e+05	2.86e+05
t-1	1.397e+05	4.39e+04	3.180	0.004	4.88e+04	2.31e+05
=====						
Omnibus:	0.888	Durbin-Watson:	1.029			
Prob(Omnibus):	0.641	Jarque-Bera (JB):	0.809			
Skew:	-0.175	Prob(JB):	0.667			
Kurtosis:	2.209	Cond. No.	4.30e+09			
=====						

GDP – ВВП

t-3 – Значения Труда с лагом t-3

t-1 – Значения Труда с лагом t-1

P>|t| – Коэффициент значимости

Adj. R-squared - Скорректированный коэффициент детерминации

Наиболее значимым оказались признаки L_{t-1} и L_{t-3}

Модель регрессии ВВП на Капитал и его лагированные значения

Полные данные по Капиталу

OLS Regression Results						
Dep. Variable:	GDP	R-squared:	0.954			
Model:	OLS	Adj. R-squared:	0.937			
Method:	Least Squares	F-statistic:	53.76			
Date:	Wed, 06 Jul 2022	Prob (F-statistic):	8.80e-11			
Time:	11:24:32	Log-Likelihood:	-704.85			
No. Observations:	26	AIC:	1426.			
Df Residuals:	18	BIC:	1436.			
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	7.131e+10	7.61e+10	0.936	0.361	-8.87e+10	2.31e+11
t-6	-1.4434	1.477	-0.977	0.341	-4.547	1.660
t-5	-0.7929	3.191	-0.248	0.807	-7.497	5.912
t-4	3.1272	3.787	0.826	0.420	-4.830	11.084
t-3	3.4292	3.589	0.955	0.352	-4.112	10.970
t-2	-2.4248	3.053	-0.794	0.437	-8.838	3.989
t-1	-5.2798	2.411	-2.190	0.042	-10.344	-0.215
t	7.2425	1.282	5.648	0.000	4.548	9.937
Omnibus:	2.249	Durbin-Watson:	1.187			
Prob(Omnibus):	0.325	Jarque-Bera (JB):	1.102			
Skew:	-0.467	Prob(JB):	0.576			
Kurtosis:	3.382	Cond. No.	1.77e+12			

GDP – ВВП

t-6 – Значения Капитала с лагом t-6

t-5 – Значения Капитала с лагом t-5

t-4 – Значения Капитала с лагом t-4

t-3 – Значения Капитала с лагом t-3

t-2 – Значения Капитала с лагом t-2

t-1 – Значения Капитала с лагом t-1

t – Значения Капитала

P>|t| – Коэффициент значимости

Adj. R-squared - Скорректированный коэффициент детерминации

Модель регрессии ВВП на Капитал и его лагированные значения

Данные после удаления наименее значимых признаков

OLS Regression Results

```
=====
Dep. Variable:          GDP      R-squared:                0.950
Model:                  OLS      Adj. R-squared:           0.941
Method:                 Least Squares  F-statistic:              100.6
Date:                   Wed, 06 Jul 2022  Prob (F-statistic):       2.20e-13
Time:                   11:27:10   Log-Likelihood:           -705.92
No. Observations:      26        AIC:                      1422.
Df Residuals:          21        BIC:                      1428.
Df Model:               4
Covariance Type:       nonrobust
=====
```

	coef	std err	t	P> t	[0.025	0.975]
const	6.265e+10	7.21e+10	0.869	0.395	-8.73e+10	2.13e+11
t-5	-4.0468	1.343	-3.013	0.007	-6.840	-1.253
t-4	6.7388	1.764	3.821	0.001	3.071	10.406
t-1	-6.1139	1.615	-3.786	0.001	-9.472	-2.756
t	7.3287	1.211	6.052	0.000	4.810	9.847

```
=====
Omnibus:                2.279   Durbin-Watson:            1.342
Prob(Omnibus):          0.320   Jarque-Bera (JB):         1.041
Skew:                   -0.420  Prob(JB):                  0.594
Kurtosis:               3.505   Cond. No.:                 1.33e+12
=====
```

GDP – ВВП

t-5 – Значения Капитала с лагом t-5

t-4 – Значения Капитала с лагом t-4

t-1 – Значения Капитала с лагом t-1

t – Значения Капитала

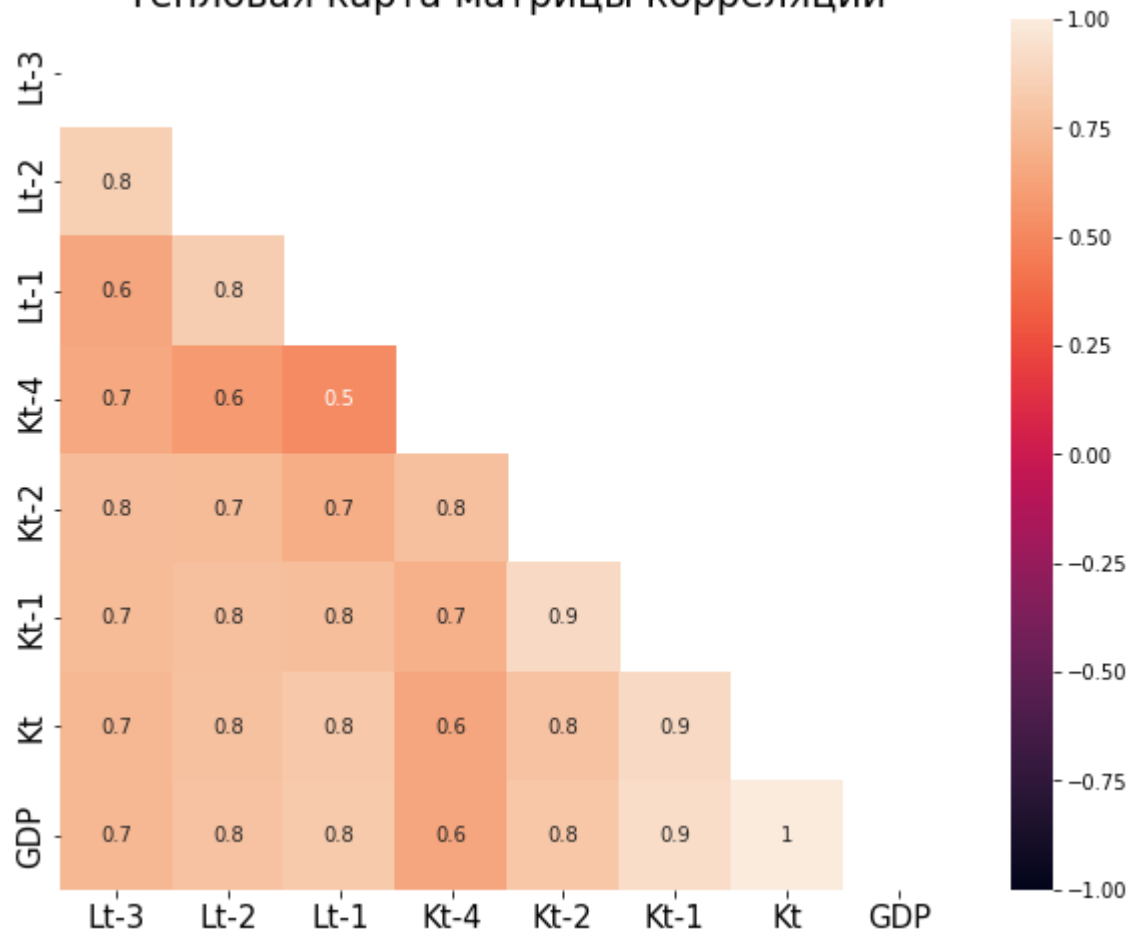
$P > |t|$ – Коэффициент значимости

Adj. R-squared - Скорректированный коэффициент детерминации

Наиболее значимым оказались признаки K_t, K_{t-1}, K_{t-4} и K_{t-5}

Наиболее коррелированные признаки

Тепловая карта матрицы корреляции



GDP – ВВП

Lt-3 – Значения Труда с лагом t-3

Lt-2 – Значения Труда с лагом t-2

Lt-1 – Значения Труда с лагом t-1

Kt-3 – Значения Капитала с лагом t-3

Kt-2 – Значения Капитала с лагом t-2

Kt-1 – Значения Капитала с лагом t-1

Kt – Значения Капитала

	Lt-3	Lt-2	Lt-1	Kt-4	Kt-2	Kt-1	Kt	GDP
Lt-3	1.000000	0.844160	0.643932	0.662092	0.754145	0.746628	0.733848	0.730726
Lt-2	0.844160	1.000000	0.837839	0.592128	0.742461	0.772577	0.780335	0.782900
Lt-1	0.643932	0.837839	1.000000	0.518391	0.686508	0.758803	0.813572	0.822484
Kt-4	0.662092	0.592128	0.518391	1.000000	0.768814	0.698830	0.641894	0.646433
Kt-2	0.754145	0.742461	0.686508	0.768814	1.000000	0.899621	0.781541	0.809733
Kt-1	0.746628	0.772577	0.758803	0.698830	0.899621	1.000000	0.901014	0.923769
Kt	0.733848	0.780335	0.813572	0.641894	0.781541	0.901014	1.000000	0.994491
GDP	0.730726	0.782900	0.822484	0.646433	0.809733	0.923769	0.994491	1.000000

Сводка по модели регрессии

$$\text{GDP} = f(\text{Lt-3}, \text{Lt-2}, \text{Lt-1}, \text{Kt-3}, \text{Kt-2}, \text{Kt-1}, \text{Kt})$$

Полные данные по модели регрессии

OLS Regression Results

Dep. Variable:	GDP	R-squared:	0.994
Model:	OLS	Adj. R-squared:	0.992
Method:	Least Squares	F-statistic:	465.5
Date:	Sun, 10 Jul 2022	Prob (F-statistic):	7.06e-20
Time:	20:25:24	Log-Likelihood:	-704.27
No. Observations:	27	AIC:	1425.
Df Residuals:	19	BIC:	1435.
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	1.713e+11	1e+12	0.170	0.866	-1.93e+12	2.27e+12
Lt-3	-8055.7606	1.31e+04	-0.613	0.547	-3.56e+04	1.95e+04
Lt-2	-6494.8459	1.63e+04	-0.399	0.694	-4.05e+04	2.75e+04
Lt-1	1.301e+04	1.38e+04	0.944	0.357	-1.58e+04	4.19e+04
Kt-4	-0.1015	0.121	-0.838	0.412	-0.355	0.152
Kt-2	0.2126	0.197	1.079	0.294	-0.200	0.625
Kt-1	0.5043	0.241	2.090	0.050	-0.001	1.009
Kt	3.5671	0.197	18.088	0.000	3.154	3.980

Omnibus:	0.177	Durbin-Watson:	1.428
Prob(Omnibus):	0.915	Jarque-Bera (JB):	0.389
Skew:	-0.038	Prob(JB):	0.823
Kurtosis:	2.417	Cond. No.	4.67e+13

GDP – ВВП

Lt-3 – Значения Труда с лагом t-3

Lt-2 – Значения Труда с лагом t-2

Lt-1 – Значения Труда с лагом t-1

Kt-3 – Значения Капитала с лагом t-3

Kt-2 – Значения Капитала с лагом t-2

Kt-1 – Значения Капитала с лагом t-1

Kt – Значения Капитала

P>|t| – Коэффициент значимости

Adj. R-squared - Скорректированный коэффициент детерминации

Сводка по модели регрессии

$GDP = f(Lt-3, Lt-2, Lt-1, Kt-3, Kt-2, Kt-1, Kt)$

Данные после удаления наименее значимых признаков

OLS Regression Results

```
=====
Dep. Variable:          GDP      R-squared:                0.993
Model:                  OLS      Adj. R-squared:           0.993
Method:                 Least Squares  F-statistic:              1726.
Date:                   Sun, 10 Jul 2022  Prob (F-statistic):      1.18e-26
Time:                   20:25:45    Log-Likelihood:          -706.63
No. Observations:      27         AIC:                    1419.
Df Residuals:          24         BIC:                    1423.
Df Model:               2
Covariance Type:       nonrobust
=====
```

```
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const      5.854e+10  2.08e+10      2.808      0.010      1.55e+10  1.02e+11
Kt-1        0.6083      0.161          3.767      0.001          0.275      0.942
Kt          3.5675      0.162          22.039     0.000          3.233      3.902
=====
```

```
=====
Omnibus:                0.822    Durbin-Watson:           1.260
Prob(Omnibus):          0.663    Jarque-Bera (JB):        0.736
Skew:                   -0.051   Prob(JB):                 0.692
Kurtosis:               2.198    Cond. No.                 7.43e+11
=====
```

GDP – ВВП

Kt-1 – Значения Капитала с лагом t-1

Kt – Значения Капитала

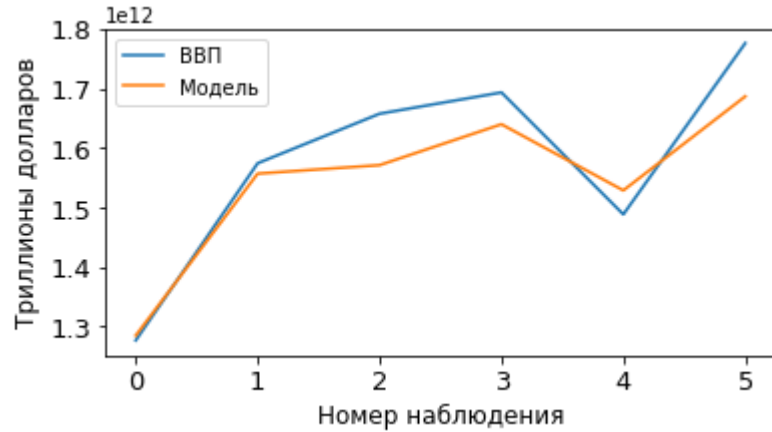
$P > |t|$ – Коэффициент значимости

Adj. R-squared - Скорректированный коэффициент детерминации

Линейная Регрессия $GDP = f(K_{t-1}, K_t)$

ВВП

Проверка модели на тестовых данных



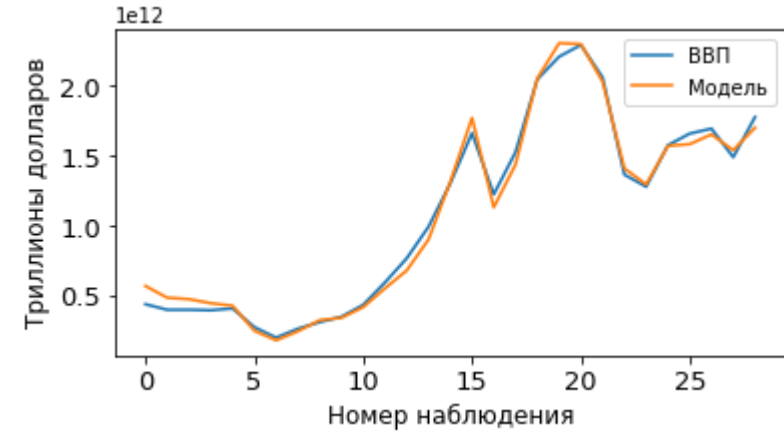
$$R^2 = 0.87$$

$$MAE = 49142314652.34\$$$

$$MSE = 3.379e+21$$

$$RMSE = 58136326816.63\$$$

Расчетные значения по модели



$$R^2 = 0.99$$

$$MAE = 49544468714.84\$$$

$$MSE = 3.786e+21$$

$$RMSE = 61534608318.31\$$$

Сравнение лучших результатов

Построение прогноза ВВП на основе производственной функции Кобба—Дугласа, где Капитал и Труд спрогнозированы на основе моделей авторегрессии

$$R^2 = 0.95$$

$$MAE = 128095480743.00\$$$

$$MSE = 2.618e+22$$

$$RMSE = 161802925312.92\$$$

Построение прогноза ВВП на основе модели Хольта-Уинтерса

$$R^2 = 0.95$$

$$MAE = 132767408003.96\$$$

$$MSE = 2.519e+22$$

$$RMSE = 158727840509.01\$$$

Построение прогноза ВВП на основе модели регрессии
 $GDP = 4.763e+10 + 0.5574 * Kt1 + 3.6528 * Kt$

$$R^2 = 0.99$$

$$MAE = 49544468714.84\$$$

$$MSE = 3.786e+21$$

$$RMSE = 61534608318.31\$$$



Выводы

1. Построен прогноз ВВП на основе производственной функции Кобба— Дугласа, где Капитал и Труд спрогнозированы на основе моделей авторегрессии.
2. Построен прогноз ВВП на основе производственной функции Кобба— Дугласа, где Капитал и Труд спрогнозированы на основе моделей модели Хольта-Уинтерса и Тейла-Вейджа.
3. Построен прогноз ВВП на основе адаптивных тренд-сезонных моделей(модели Хольта-Уинтерса и Тейла-Вейджа).
4. Построен прогноз ВВП на основе модели регрессии $GDP = f(Lt-3, Lt-2, Lt-1, Kt-3, Kt-2, Kt-1, Kt)$
5. Наиболее точный результат($R^2 = 0.99$) показала модель регрессии следующего вида:

$$GDP = 4.763e+10 + 0.5574 * K_{t1} + 3.6528 * K_t$$

