

Львівський національний університет імені Івана Франка

Економічний факультет

Кафедра статистики

ЛАБОРАТОРНИЙ ПРАКТИКУМ

з курсу «СТАТИСТИЧНИЙ АНАЛІЗ ЧАСОВИХ

РЯДІВ»

З ВИКОРИСТАННЯМ EXCEL ТА STATISTICA

Львів-2022

До друку рекомендовано кафедрою статистики

Львівського національного університету імені Івана Франка (протокол №_1_
від 30 серпня 2022 р.)

Автор-укладач:

Вільчинська Оксана Миколаївна, Львівський національний університет, доцент
кафедри статистики.

Рецензенти:

доктор наук з державного управління,
професор Мамонова Валентина Василівна,
професор кафедри менеджменту і публічного адміністрування
Харківського національного університету міського господарства
імені О.М. Бекетова.

Вдовин Мар'яна Любомирівна,
Львівський національний університет,
доцент кафедри статистики.

Відповідальний за випуск:

Семен Олексійович Матковський,
професор, завідувач кафедри статистики
Львівського національного університету;

ЗМІСТ

Завдання на лабораторні роботи для аналізу часових рядів в EXCEL.....	4
Методичні поради	4
1. Розрахунок показників аналізу рядів динаміки	4
2. Визначення параметрів лінійного рівняння та оцінка якості тренду	8
3. Аналітичне вирівнювання за рівнянням гіперболи, параболи другого порядку, експоненти.....	15
4. Розрахунок індексів сезонності	17
Завдання для лабораторних робіт з побудови моделі декомпозиції часових рядів у ПП STATISTICA.....	21
Методичні поради	21

Завдання для лабораторних робіт з аналізу часових рядів в EXCEL

1. На основі щомісячних даних сформувати ряд динаміки та розрахувати:

→ Базові та ланцюгові абсолютні прирости, темпи зростання і приросту
Абсолютне значення 1% приросту (ланцюгового).

→ Середній рівень ряду.

→ Середній абсолютний приріст (двома способами).

→ Середній темп зростання і приросту.

→ Визначити трендові рівняння за рівнянням прямої лінії та здійснити точковий та інтервальний прогноз на наступний рік з ймовірністю 0,95 (Для виконання завдання ймовірність $p=1 - 0,01*k$).

→ Визначити трендові рівняння за рівнянням параболи другого порядку, експоненти.

2. На основі квартальних даних за три роки розрахувати індекси сезонності методом аналітичного вирівнювання за рівнянням прямої лінії та оцінити інтенсивність сезонності.

Методичні поради

1. Розрахунок показників аналізу рядів динаміки

Вихідні дані до задачі показано на рис.73

6	січень	17,0	24,7	18,4	19,4	23,3	23,7	16,4
7	лютий	22,8	20,9	21,6	15,4	21,1	21,0	22,8
8	березень	18,3	23,0	24,0	23,4	14,6	21,9	18,3
9	квітень	20,1	21,2	19,1	19,1	16,8	16,8	20,1
10	травень	24,4	12,9	21,8	23,7	24,8	17,7	24,4
11	червень	16,7	24,9	21,0	19,6	16,2	17,8	16,7
12	липень	24,8	20,8	13,5	24,0	14,2	24,3	24,8
13	серпень	15,5	19,5	21,7	15,7	14,1	17,6	15,5
14	вересень	18,3	23,6	15,4	13,5	14,3	13,4	18,3
15	жовтень	15,3	20,8	19,2	16,3	15,4	25,0	16,2
16	листопад	17,2	22,9	24,7	19,9	16,4	23,8	17,2
17	грудень	15,8	19,5	23,7	19,3	15,7	21,2	15,8
18	Всього	226,2	254,8	244,4	229,2	206,8	244,3	226,5
19								

Рис. 1. Приклад робочого вікна з вихідними даними для виконання роботи

Оскільки, рівні даного ряду є абсолютними аддитивними величинами, то:

1. Середньорічний рівень ряду розраховуємо за формулою середньої арифметичної простої:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n},$$

де y_i - рівні інтервального ряду динаміки, n – кількість рівнів ряду.

2. Абсолютні прирости обчислюємо як різницю двох рівнів ряду:

базисні абсолютні прирости: $\Delta_i^b = y_i - y_1$

ланцюгові абсолютні прирости: $\Delta_i^l = y_i - y_{i-1}$,

де i – порядковий номер рівня ряду, y_i – поточний рівень ряду, y_{i-1} , рівень ряду, який передує y_i , y_1 – базисний рівень ($i \in [1, n]$).

3. Середній абсолютний приріст можемо розраховувати як:

– відношення суми абсолютних приростів до їхньої кількості: $\bar{\Delta} = \frac{\sum \Delta}{n-1}$;

– різниці поточного та початкового рівнів ряду до їхньої кількості зменшеної на 1: $\bar{\Delta} = \frac{y_n - y_0}{n-1}$.

4. Коефіцієнти зростання обчислюємо як відношення відповідних рівнів ряду:

базисний коефіцієнт зростання $k_o^z = \frac{y_i}{y_1}$

ланцюговий коефіцієнт зростання $k_n^z = \frac{y_i}{y_{i-1}}$

Темпи зростання – як добуток відповідних коефіцієнтів зростання на 100%:

базисний темп зростання $T_o^z = k_o^z \cdot 100$

ланцюговий темп зростання $T_n^z = k_n^z \cdot 100$

Темпи приросту – як різницю відповідного темпу зростання і 100%:

базисний темп приросту $T_o^{IP} = T_o^z - 100\%$

ланцюговий темп приросту $T_n^{IP} = T_n^z - 100\%$

Середньорічний темп зростання та приросту розраховуємо за середньою геометричною:

$$T^z = \sqrt[n-1]{T_1^z \cdot \dots \cdot T_{n-1}^z} \text{ або } T^z = \sqrt[n-1]{\frac{y_n}{y_1}}.$$

Вихідні дані в нас містять показники за відповідні місяці та за роки. При розв'язанні першої задачі використовуємо дані за відповідні роки. Для полегшення розрахунків їх можна представити у зручнішому вигляді (роки і відповідні їм дані у стовбцях). зручно це зробити наступним способом: скопіювати відповідні дані і вставити через закладку *Спеціальная вставка* —> *Значения +Транспонировать*. Практична реалізація даного моменту зображена на рис. 74.

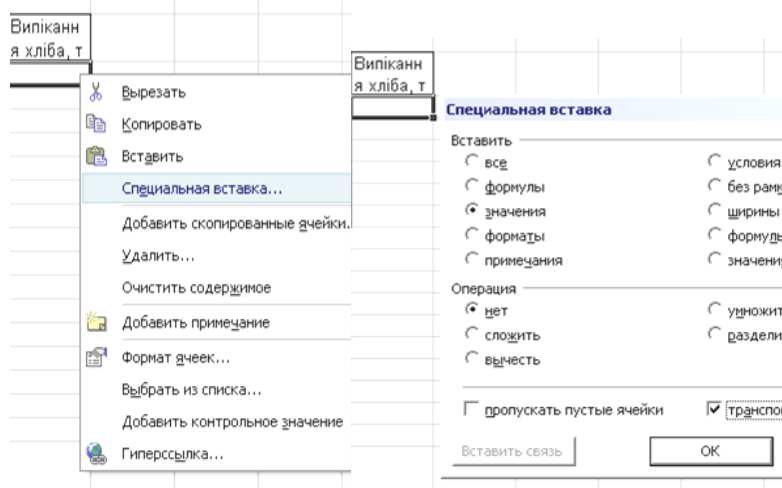


Рис. 2. Порядок застосування опції „Спеціальна вставка”

Застосувавши описані вище формули та команди обчислюємо потрібні показники. Відповідні команди, формули та результати їх застосування показано на рис. 75, 76.

=ОКРУГЛ(СУММ(В28:В34)/СЧЕТ(В28:В34),2)										
В	С	Д	Е	Ф	Г	Н	І	Ј	К	
На основі даних ряду динаміки розрахувати:										
1. Середньорічне випікання хліба										
=ОКРУГЛ(СУММ(В28:В34)/СЧЕТ(В28:В34),2)										
іві та ланцюгові абсолютні прирости, темпи зростання та приросту, абсолютне значення 1% приросту										
Випікання хліба, т	Абсолютні прирости, т		Темп зростання, %		Темп приросту, %		Абсолютне значення 1% приросту		кофіцієнт зростання	
	базовий	ланцюговий	базовий	ланцюговий	базовий	ланцюговий				
	$\Delta_i^a = y_i - y_1$	$\Delta_i^l = y_i - y_{i-1}$	$k_i^a = \frac{y_i}{y_1} \cdot 100$	$k_i^l = \frac{y_i}{y_{i-1}} \cdot 100$	$T_i^a = \frac{y_i}{y_1} - 100$	$T_i^l = \frac{y_i}{y_{i-1}} - 100$			A1%	
220,1	—	—	—	—	—	—	—	—	—	—
254,8	=B29-B\$28	=B29-B28	=ОКРУГЛ(В29/В\$28*100,2)	=ОКРУГЛ(В29/В28*100,2)	=E29-100	=F29-100	=D29/Н29	=B28/100	=F29*0,01	
244,4	=B30-B\$28	=B30-B29	=ОКРУГЛ(В30/В\$28*100,2)	=ОКРУГЛ(В30/В29*100,2)	=E30-100	=F30-100	=D30/Н30	=B29/100	=F30*0,01	
218,2	=B31-B\$28	=B31-B30	=ОКРУГЛ(В31/В\$28*100,2)	=ОКРУГЛ(В31/В30*100,2)	=E31-100	=F31-100	=D31/Н31	=B30/100	=F31*0,01	
206,8	=B32-B\$28	=B32-B31	=ОКРУГЛ(В32/В\$28*100,2)	=ОКРУГЛ(В32/В31*100,2)	=E32-100	=F32-100	=D32/Н32	=B31/100	=F32*0,01	
244,3	=B33-B\$28	=B33-B32	=ОКРУГЛ(В33/В\$28*100,2)	=ОКРУГЛ(В33/В32*100,2)	=E33-100	=F33-100	=D33/Н33	=B32/100	=F33*0,01	
226,5	=B34-B\$28	=B34-B33	=ОКРУГЛ(В34/В\$28*100,2)	=ОКРУГЛ(В34/В33*100,2)	=E34-100	=F34-100	=D34/Н34	=B33/100	=F34*0,01	
зній абсолютний приріст можна розрахувати										
як відношення суми			різниця n-ного та 1 рівня ряду							
=СУММ(D29:D34)/6			=(B34-B28)/6							
дньорічний темп зростання та приросту розраховуємо за середньою геометричною										
=СТЕПЕНЬ(В34/В28,1/6)*100			=СТЕПЕНЬ((К29*К30*К31*К32*К33*К34,1/6)*100							

Рис. 3. Опис формул для обчислення показників динаміки

4. Середньорічний темп зростання та приросту розраховуємо за середньою геометричною											
В		C		D		E		F		G	
нові дані ряду динаміки розрахувати:											
середньорічне випікання хліба											
230,73											
потні прирости, темпи зростання та приросту, абсолютне значення											
Випікання хліба, т	Абсолютні прирости, т		Темп зростання, %		Темп приросту, %		Абсолютне значення 1 % приросту		кофіцієнт зростання		
	базовий	ланцюговий	базовий	ланцюговий	базовий	ланцюговий					
y_i	$\Delta_i^b = y_i - y_1$	$\Delta_i^c = y_i - y_{i-1}$	$k_i^b \cdot 100$	$k_i^c \cdot 100$	$T_i^b - 100$	$T_i^c - 100$	A1%				
220,1	—	—	—	—	—	—	—		—		
254,8	34,7	34,7	115,77	115,77	15,77	15,77	2,20	2,20	1,1577		
244,4	24,3	-10,4	111,04	95,92	11,04	-4,08	2,55	2,55	0,9592		
218,2	-1,9	-26,2	99,14	89,28	-0,86	-10,72	2,44	2,44	0,8928		
206,8	-13,3	-11,4	93,96	94,78	-6,04	-5,22	2,18	2,18	0,9478		
244,3	24,2	37,5	111,00	118,13	11,00	18,13	2,07	2,07	1,1813		
226,5	6,4	-17,8	102,91	92,71	2,91	-7,29	2,44	2,44	0,9271		
най абсолютний приріст можна розраховувати як відношення суми абсолютних різниць n-ного та 1 рівня ряду до кількості рівнів ряду мінус 1											
1,066667			1,066667								
середньорічний темп зростання та приросту розраховуємо за середньою геометричною											
100,4788588											
100,4794261											

Рис. 4. Результат застосування формул для обчислення показників динаміки

2. Визначення параметрів лінійного рівняння та оцінка якості тренду

Вихідні дані до задачі містяться на рис.77. Представимо даний динамічний ряд графічно (рис. 73).

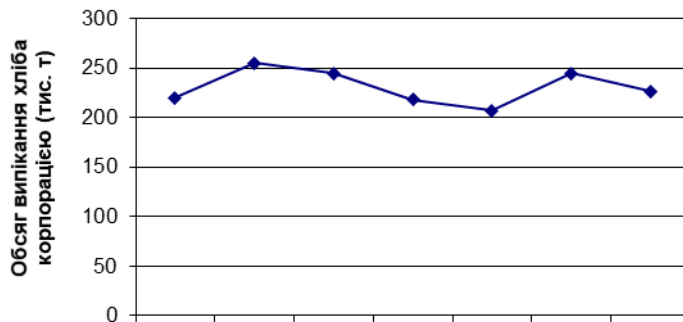


Рис. 5. Динаміка випікання хліба корпорацією.

Нагадаємо, що суть аналітичного вирівнювання динамічних рядів полягає в тому, що фактичні рівні ряду замінюють плавними рівнями, які обчислюють на основі певної прямої чи кривої, обраної з припущення, що вона найточніше відображає загальну тенденцію явища.

Щоб показати як спростити розрахунки (так зване використання EXCEL у ролі калькулятора) перенесемо початок відліку часу в середину ряду. Для

непарної кількості рівнів динамічного ряду це роблять так, як показано в розрахунковій таблиці:

Роки	y_i	t_i	$y_i t_i$
2014	220,1	-3	-660,3
2015	254,8	-2	-509,6
2016	244,4	-1	-244,4
2017	218,2	0	0
2018	206,8	1	206,8
2019	244,3	2	488,6
2020	226,5	3	679,5
Всього	1615,1	0	-718,9

Для визначення параметрів лінійної функції скористаємося відповідною системою рівнянь:

$$a = \frac{\sum y_i}{n} = \frac{1615,1}{7} = 230,73;$$

$$b = \frac{\sum y_i t_i}{\sum t_i^2} = -1,407.$$

Модель трендового рівняння: $y(t) = 230,73 - 1,407 \cdot t$.

Для точкового прогнозу на найближчі роки (за умови, що така тенденція розвитку явища збережеться в майбутньому) необхідно у рівняння тренду підставити відповідне значення параметра t . Наприклад, для 2021 року

$$2021 \text{ рік: } y(4) = 230,73 - 1,407 \cdot t = 230,73 - 1,407 \cdot 4 = 225,1.$$

Для інтервального прогнозу будемо довірчий інтервал, для якого необхідно розрахувати стандартну похибку апроксимації:

$$S_t = \sqrt{\frac{\sum_{i=1}^n (y_i - y(t))^2}{n-l}} = 18,7.$$

Для визначення критерію Стюдента з таблиць треба визначити рівень істотності та кількість ступенів довіри:

рівень істотності: $\alpha = 0,05$;

кількість ступенів довіри: $df = 7 - 1 = 6$;

$t_\alpha = 2,447$;

$t_\alpha \cdot S_t = 2,447 \cdot 18,7 = 45,9$.

Інтервальний прогноз:

2021 рік: $225,1 \pm 45,9$.

Отже, якщо дана тенденція збережеться, з імовірністю 95% у 2021 р. корпорація може випустити від 179,2 до 271,0 т хліба.

У програмі Microsoft Excel прогнозування можна здійснювати без перетворення параметра часу. Ми розглянемо два способи для побудови лінійного тренду.

Найпоширенішим способом прогнозування на основі лінійного тренду та оцінки його якості за допомогою програми Microsoft Excel є використання функції *TREND* / *ТЕНДЕНЦИЯ* (відомі_значення_y; відомі_значення_x; нові_значення_x; конст). Побудову лінійного тренду, точковий, інтервальний прогнози та розрахунок показників оцінки якості обчислених показників показано на рис. 78.

fx = ОКРУГЛ(СРЗНАЧ(С3:С9);2)			
С		D	E
Випікання хліба, т			
yi	t	y(t)	
220,1	-3	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;;1)	
254,8	-2	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;;1)	
244,4	-1	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;;1)	
218,2	0	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;;1)	
206,8	1	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;;1)	
244,3	2	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;;1)	
226,5	3	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;;1)	
y=230,73-1,407t			

11			
12	Точковий прогноз		
13		=D9+1	=ТЕНДЕНЦИЯ(С3:С9;В3:В9;В13;1)
14	Інтервальний прогноз		
15		=E13-C\$22*C\$23	=E13+C\$22*C\$23
16	Оцінка якості		
17	Опис	Формула	
18	n	=СЧЕТ(С2:С9)	
19	a	=ОКРУГЛ(СРЗНАЧ(С3:С9);2)	
20	b	=ОКРУГЛ(НАКЛОН(С3:С9;D3:D9);3)	
21	R ²	=КВАДРОТКЛ(Е3:Е9)/КВАДРОТКЛ(С3:С9)	
22	S _t	=ОКРУГЛ(СТОШУХ(С3:С9;D3:D9);2)	
23	t	=ОКРУГЛ(СТЫЮДРАСПОБР(0,05;С18-1);2)	
24			

Рис. 6. Застосування функції *TREND* / *ТЕНДЕНЦИЯ* до побудови лінійного тренду

Результат застосування функції *TREND* / *ТЕНДЕНЦИЯ* (відомі_значення_y; відомі_значення_x; нові_значення_x; конст) показано на Рис. 7. Як видно з порівняння результатів розрахунків здійснених класичним

способом та за допомогою програми Microsoft Excel з використанням функції *TREND / ТЕНДЕНЦИЯ* практично не відрізняються.

В	С	Д	Е	F
Рік	Випікання хліба, т			
	y_i	t	$y(t)$	
2002	220,1	-3	234,95	
2003	254,8	-2	233,5428571	
2004	244,4	-1	232,1357143	
2005	218,2	0	230,7285714	
2006	206,8	1	229,3214286	
2007	244,3	2	227,9142857	
2008	226,5	3	226,5071429	
$y=230,73-1,407t$				

Інтервальний прогноз	
	179,187
	271,013
Оцінка якості побудованого тренду	
Опис	результат
n	7
a	230,73
b	-1,407
R^2	0,030615159
S_t	18,74
t	2,45

Рис. 7. Результат застосування функції *TREND / ТЕНДЕНЦИЯ* до побудови лінійного тренду.

Також для побудови лінійного тренду за допомогою програми Microsoft Excel можна використовувати функцію *LINEST /ЛИНЕЙН* (відомі_значення_y; відомі_значення_x; конст; статистика). Потрібно зауважити, що *LINEST /ЛИНЕЙН* – функція масиву, і для того щоб її правильно ввести, спочатку потрібно виділити діапазон в якому хочемо мати виведені обчислені результати, викликати функцію (заповнення відповідних полів показано на Рис. 8) та ввести натисненням комбінації клавіш CTRL+SHIFT+ENTER.

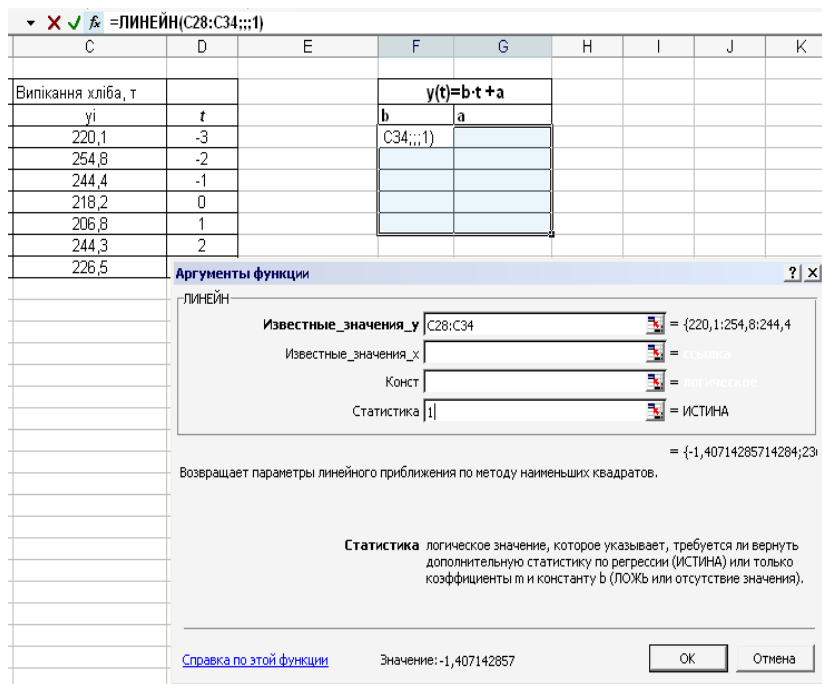


Рис. 8. Діалогове вікно функції *LINEST /ЛИНЕЙН*

Результат застосування функції *LINEST /ЛИНЕЙН* показано на Рис. 9. Ми отримали рівняння: $y(t) = 236,4 - 1,407t$; похибка параметра b: **3,54**; похибка параметра a: **15,84**; стандартна похибка: **18,74**; коефіцієнт детермінації: **0,0306**.

Випікання хліба, т		y(t)=b·t+a	
y _i	t	b	a
220,1	-3	-1,407143	236,3571429
254,8	-2	3,541058	15,83609285
244,4	-1	0,0306152	18,73751775
218,2	0	0,1579102	5
206,8	1	55,441429	1755,472857
244,3	2		
226,5	3		

Рис. 9. Результат застосування функції *LINEST /ЛИНЕЙН*

Як видно з Рис. 9 та Рис. 7, основні характеристики, розраховані за допомогою функції *LINEST /ЛИНЕЙН* практично не відрізняються від попередніх способів.

Алгоритм графічного представлення показано на Рис. 10, Рис. 11, Рис. 12, Рис. 13. Практична реалізація: за допомогою *Мастер діаграм* побудувати графік → виділити лінію динамічного ряду і в меню правої клавіші миші вибрати *Додати лінію тренду* → у вкладці *Тип* вибираємо *Лінійна* → у вкладці *параметри* вибираємо *показувати рівняння на діаграмі* і *показувати величину вірогідності апроксимації*.

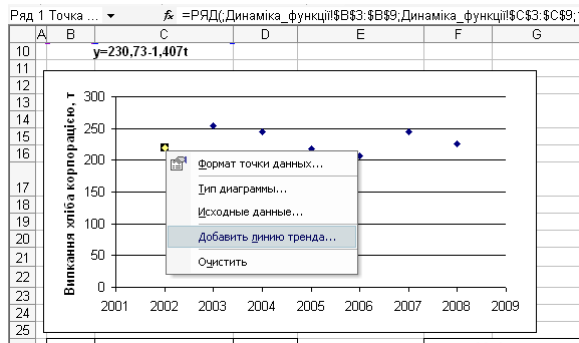


Рис. 10

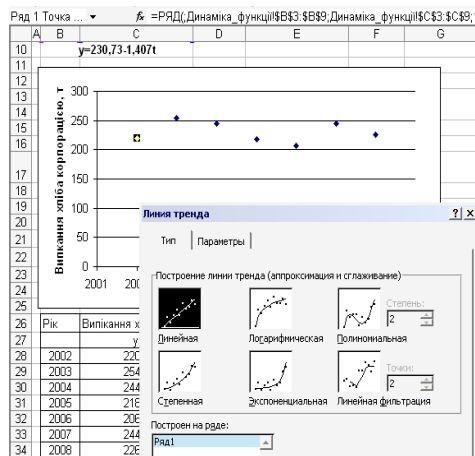


Рис. 11

Линия тренда [?] [X]

Тип | Параметры

Название аппроксимирующей (сглаженной) кривой

автоматическое: **Линейный (Ряд1)**

другое: _____

Прогноз

вперед на: единиц

назад на: единиц

пересечение кривой с осью Y в точке:

показывать уравнение на диаграмме

поместить на диаграмму величину достоверности аппроксимации (R^2)

Рис. 12

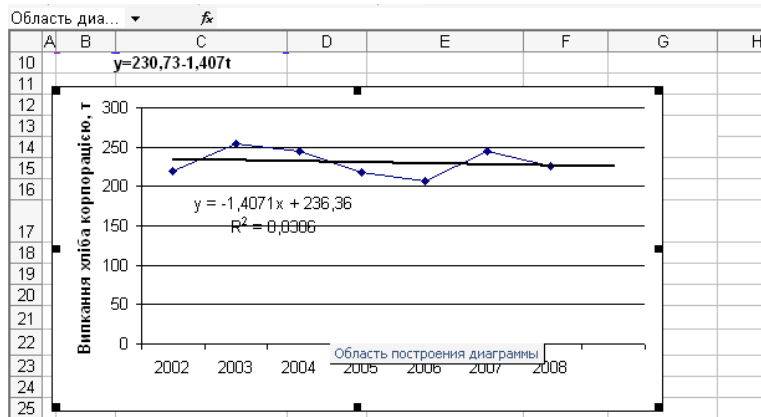


Рис. 13

3. Аналітичне вирівнювання за рівнянням гіперболи, параболи другого порядку, експоненти

У випадку аналітичного вирівнювання ряду динаміки за допомогою гіперболи, потрібно звести нелінійне рівняння гіперболи до лінійного вигляду, тобто рівняння гіперболи $y(t) = a + \frac{b}{t}$ шляхом заміни змінних $z = \frac{1}{t}$ перетворюється у рівняння вигляду: $y^*(t) = a + bz$. Всі подальші розрахунки проводимо як для лінійного випадку.

Аналітичне вирівнювання за допомогою параболи другого порядку проводимо аналогічно до рівняння прямої, за допомогою функції *LINEST* /*ЛИНЕЙН* (відомі_значення_y; відомі_значення_x; конст; статистика), з відмінністю, що в полі *відомі_значення_x* вводимо діапазон, де розміщені значення t та t^2 . Практична реалізація зображена на рис.14

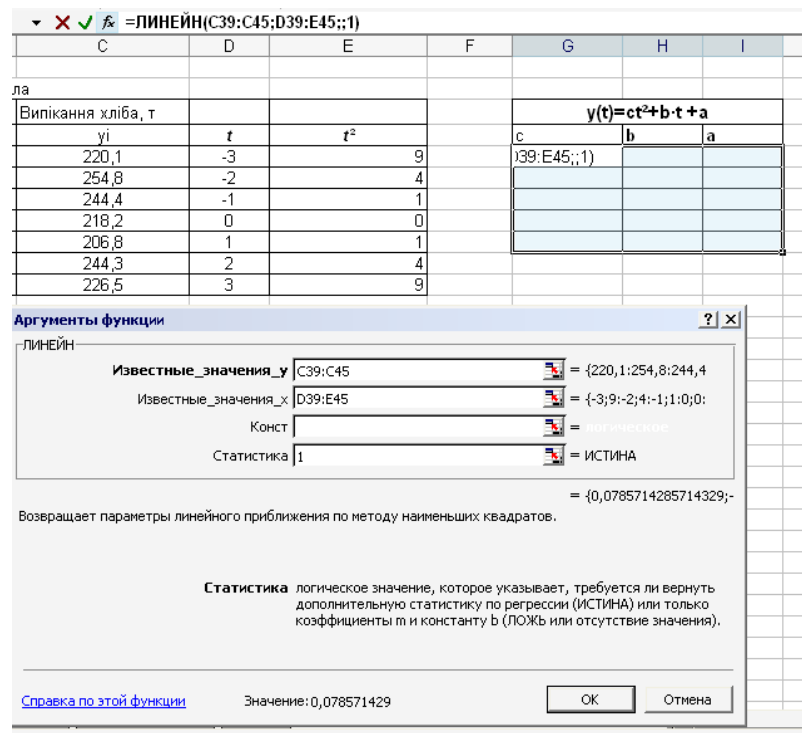


Рис. 14. Діалогове вікно функції *LINEST* /*ЛИНЕЙН* для побудови параболи другого порядку

Шукане рівняння параболи та оцінка його параметрів показані на Рис. 15. Аналогічно як і у випадку прямої лінії в клітинках *G40:I40* виведені похибки

вдповідних параметрів, $G41$ – коефіцієнт детермінації, $H41$ – стандартна похибка.

y(t)=0,0786t ² -1,407t+230,4						
C	D	E	F	G	H	I
a						
Випікання хліба, т			y(t)=ct ² +b·t+a			
	t	t ²	c	b	a	
220,1	-3	9	0,078571429	-1,40714	230,4143	
254,8	-2	4	2,285405485	3,958438	12,09323	
244,4	-1	1	0,030901518	20,94609	#N/D	
218,2	0	0	0,063773741	4	#N/D	
206,8	1	1	55,96	1754,954	#N/D	
244,3	2	4				
226,5	3	9				
y(t)=0,0786t ² -1,407t+230,4						

Рис. 15. Рівняння параболи та оцінка його параметрів.

Графічне представлення здійснюється за тим же алгоритмом, що у випадку прямої лінії, з відмінністю, що у вкладці *Тип* вибираємо *Полиномиальная* (у полі для встановлення кількості змінних – 2). (Рис. 16)

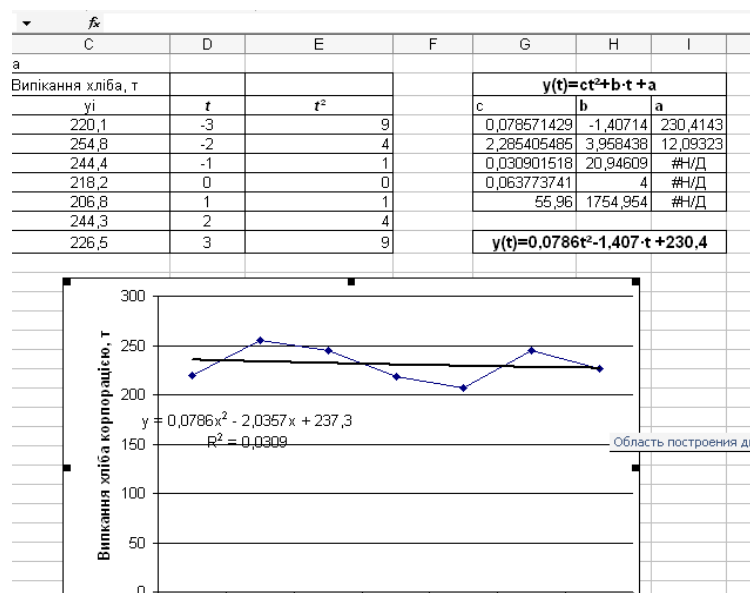


Рис. 16. Результат графічного представлення аналітичного вирівнювання за парабою другого порядку

Розрахунок параметрів показникового рівняння у програмі Microsoft Excel проводимо за допомогою функції *LOGEST* / *ЛГРФПРИБЛ* (відомі_значення_y; відомі_значення_x; конст; статистика). *LOGEST* / *ЛГРФПРИБЛ* – функція масиву. Для виведення її результатів виділяємо діапазон, і заповнюємо відповідні поля в діалоговому вікні (Рис. 17) та вводимо натисненням комбінації клавіш *CTRL+SHIFT+ENTER*.

Шукане рівняння параболи та оцінка його параметрів показані на Рис. 18. Аналогічно як і у випадку прямої лінії в клітинках G52: H 52 виведені похибки відповідних параметрів, G53 – коефіцієнт детермінації, H53 – стандартна похибка.

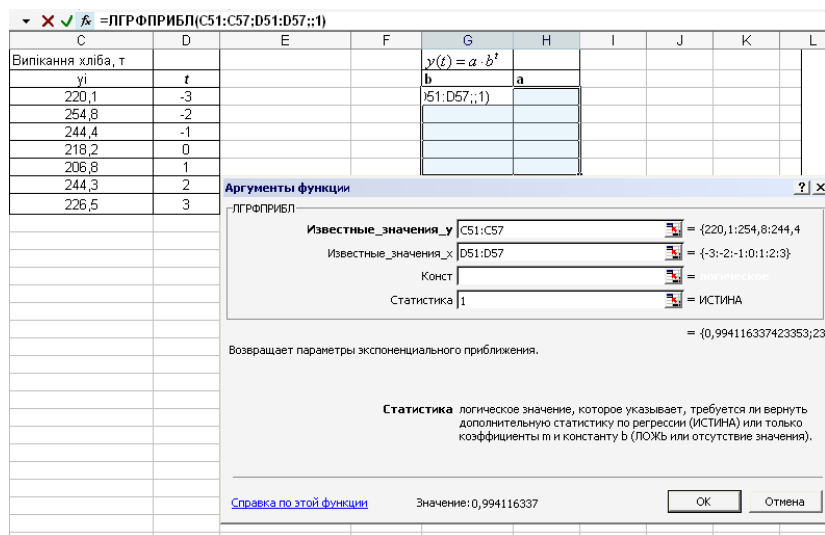


Рис. 17. Діалогове вікно функції LOGEST / ЛГРФПРИБЛ

Випікання хліба, т	t	$y(t) = a \cdot b^t$	a
220,1	-3	0,994116337	230,1682
254,8	-2	0,01537231	0,030745
244,4	-1	0,028628216	0,081343
218,2	0	0,147359725	5
206,8	1	0,000975023	0,033083
244,3	2		
226,5	3		
		$y=230,12 \cdot 0,994^t$	

Рис. 18. Результат застосування функції LOGEST / ЛГРФПРИБЛ

Висновок. Якщо рівняння тренду підібрано вдало, то йому мають бути притаманними близькі до нуля похибки (параметрів і рівняння загалом) та коефіцієнт детермінації, близький до 1.

4. Розрахунок індексів сезонності

На основі місячних даних розрахувати індекси сезонності методом аналітичного вирівнювання за рівнянням прямої лінії та оцінити інтенсивність сезонності.

Ряди динаміки можуть бути представлені у вигляді суми компонент: основної тенденції розвитку – *тренду*, *сезонної* (періодичної складової та

випадкової компоненти. приклади побудови тренду наведені в задачі 1. Тут ми зупинимось на сезонній складовій, котра має важливе практичне значення.

В статистиці існує ряд методів вивчення та виміру сезонних коливань:

- метод абсолютних різниць;
- метод відносних різниць;
- побудова індексів сезонності;
- побудова аналітичної моделі.

Наше завдання побудувати індекси сезонності та оцінити інтенсивність сезонності. Вихідні дані до задачі та розрахунок індексів сезонності показано на Рис. 19 та Рис. 21

95	січень	17,0	24,7	18,4	19,4	23,3	23,7	16,4
96	лютий	22,8	20,9	21,6	15,4	21,1	21,0	22,8
97	березень	18,3	23,0	24,0	23,4	14,6	21,9	18,3
98	квітень	20,1	21,2	19,1	19,1	16,8	16,8	20,1
99	травень	24,4	12,9	21,8	23,7	24,8	17,7	24,4
100	червень	16,7	24,9	21,0	19,6	16,2	17,8	16,7
101	липень	24,8	20,8	13,5	24,0	14,2	24,3	24,8
102	серпень	15,5	19,5	21,7	15,7	14,1	17,6	15,5
103	вересень	18,3	23,6	15,4	13,5	14,3	13,4	18,3
104	жовтень	15,3	20,8	19,2	16,3	15,4	25,0	16,2
105	листопад	17,2	22,9	24,7	19,9	16,4	23,8	17,2
106	грудень	15,8	19,5	23,7	19,3	15,7	21,2	15,8
107	Всього	226,2	254,8	244,4	229,2	206,8	244,3	226,5
108								
109								

Рис. 19. Вихідні дані до задачі

B109		fx =СУММ(B107:H107)/(12*7)	
	A	B	C
108			
109	Середньомісячне випікання		
110	хліба за сім років=		
111		=СУММ(B107:H107)/(12*7)	
	середньомісячне випікання	Індекс сезонності	
112	хліба		
113		коефіцієнт	%
114	=ОКРУГЛ(СУММ(B95:H95)/7,2)	=ОКРУГЛ(A114/B\$109;4)	=B114*100
115	=ОКРУГЛ(СУММ(B96:H96)/7,2)	=ОКРУГЛ(A115/B\$109;4)	=B115*100
116	=ОКРУГЛ(СУММ(B97:H97)/7,2)	=ОКРУГЛ(A116/B\$109;4)	=B116*100
117	=ОКРУГЛ(СУММ(B98:H98)/7,2)	=ОКРУГЛ(A117/B\$109;4)	=B117*100
118	=ОКРУГЛ(СУММ(B99:H99)/7,2)	=ОКРУГЛ(A118/B\$109;4)	=B118*100
119	=ОКРУГЛ(СУММ(B100:H100)/7,2)	=ОКРУГЛ(A119/B\$109;4)	=B119*100
120	=ОКРУГЛ(СУММ(B101:H101)/7,2)	=ОКРУГЛ(A120/B\$109;4)	=B120*100
121	=ОКРУГЛ(СУММ(B102:H102)/7,2)	=ОКРУГЛ(A121/B\$109;4)	=B121*100
122	=ОКРУГЛ(СУММ(B103:H103)/7,2)	=ОКРУГЛ(A122/B\$109;4)	=B122*100
123	=ОКРУГЛ(СУММ(B104:H104)/7,2)	=ОКРУГЛ(A123/B\$109;4)	=B123*100
124	=ОКРУГЛ(СУММ(B105:H105)/7,2)	=ОКРУГЛ(A124/B\$109;4)	=B124*100
125	=ОКРУГЛ(СУММ(B106:H106)/7,2)	=ОКРУГЛ(A125/B\$109;4)	=B125*100
126			

Рис. 20. Розрахунок індексів сезонності (формули)

A114		fx = ОКРУГЛ(СУММ(B95:H95)/7		
	A	B	C	D
108				
109	Середньомісяч	19,43095238		
110	не випікання			
111	хліба за сім			
112	середньомісяч	Індекс сезонності		
	не випікання			
	хліба			
113		коефіцієнт	%	
114	20,43	1,0514	105,14	
115	20,82	1,0715	107,15	
116	20,5	1,055	105,5	
117	19,04	0,9799	97,99	
118	21,39	1,1008	110,08	
119	18,99	0,9773	97,73	
120	20,89	1,0751	107,51	
121	17,07	0,8785	87,85	
122	16,7	0,8595	85,95	
123	18,31	0,9423	94,23	
124	20,32	1,0458	104,58	
125	18,71	0,9629	96,29	
126				

Рис. 21 Розрахунок індексів сезонності (результат).

Лінія тренду, яка описує щомісячну динаміку обсягів випікання хліба протягом семи років має вигляд $y(t) = -0,018 \cdot t + 20,2$. На Рис. 22 наведені теоретичні рівні ряду динаміки, розраховані на підставі побудованої лінії тренду, а також показники тренду, скоригованого на сезонність.

fx = D245*B114					
B	C	D	E	F	G
Місяць, t	Випікання хліба	Тренд		Тренд, скоригований на сезонність	
	y(t)	y(t) = -0,018·t + 20,2		y(t)·Iсез	
1	17,0	20,182		21,21935	
63	21,9	19,066		20,11463	
64	16,8	19,048		18,66514	
65	17,7	19,03		20,94822	
66	17,8	19,012		18,58043	
67	24,3	18,994		20,42045	
68	17,6	18,976		16,67042	
69	13,4	18,958		16,2944	
70	25,0	18,94		17,84716	
71	23,8	18,922		19,78863	
72	21,2	18,904		18,20266	
73	16,4	18,886		19,85674	
74	22,8	18,868		20,21706	
75	18,3	18,85		19,88675	
76	20,1	18,832		18,45348	
77	24,4	18,814		20,71045	
78	16,7	18,796		18,36933	
79	24,8	18,778		20,18823	
80	15,5	18,76		16,48066	
81	18,3	18,742		16,10875	
82	16,2	18,724		17,64363	
83	17,2	18,706		19,56273	
84	15,8	18,688		17,99468	

Рис. 22. Тренд і сезонні коливання виробництва хліба.

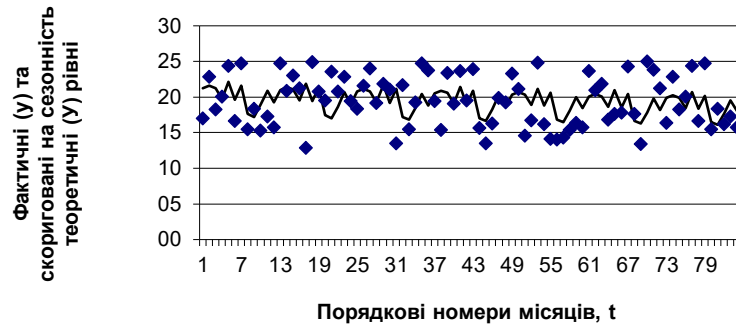


Рис. 23. Динаміка вирибництва хліба корпорацією за 2014-2020 рр. та скоригований на сезонність лінійний тренд динаміки

Загальним показником сили коливання динамічного ряду сезонності за рік є середнє квадратичне відхилення індексів сезонності представлене у відсотках:

$$\sigma_{s_i} = \sqrt{\frac{\sum (I_{s_i} - 1)^2}{12}} 100\% = 7,57$$

Чим меншим буде обчислений показник, тим меншою буде сезонність явища, котре досліджуємо. В нашому випадку можна стреджувати, що сезонність у динаміці випікання хліба є дуже малою.

Завдання для лабораторних робіт з побудови моделі декомпозиції часових рядів у ПП STATISTICA

1. Визначити вид моделі декомпозиції.
2. Виявити всі складові.
3. Спрогнозувати трендову складову.
4. Здійснити спектральний аналіз циклічної складової.
5. Здійснити композицію моделі та перевірити її якість.
6. Сформулювати висновки.

Методичні поради

Необхідно сформувати динамічний ряд і подати його у вигляді файла в ППП Statistica. Маємо дані про обсяг імпорту природного газу США за 1973-2017 рр. (<https://www.eia.gov/dnav/ng/hist/n9100us2m.htm>), які наведені на рис. 1.

	1 T	2 Gaz
1	Jan-1973	92694
2	Feb-1973	83870
3	Mar-1973	91581
4	Apr-1973	88407
5	May-1973	85844
6	Jun-1973	79121
7	Jul-1973	79428
8	Aug-1973	84400
9	Sep-1973	81157
10	Oct-1973	92151
11	Nov-1973	85784
12	Dec-1973	88466
13	Jan-1974	85815
14	Feb-1974	78919
15	Mar-1974	85190
16	Apr-1974	83371
17	May-1974	79764
18	Jun-1974	74245
19	Jul-1974	73247
20	Aug-1974	76037
21	Sep-1974	70511
22	Oct-1974	82625
23	Nov-1974	83107
24	Dec-1974	86454
25	Jan-1975	80824
26	Feb-1975	74496
27	Mar-1975	83380
28	Apr-1975	82132

Рисунок 1. Фрагмент вихідних даних

Наше завдання виявити основні закономірності, виокреми ти основні компоненти та здійснити прогноз на певний період. Оцінювати якість прогнозу будемо за допомогою застосування процедури крос-перевірки.

За допомогою автокореляційного аналізу визначимо наявність сезонності та ширину сезонного лагу.

Працювати будемо в модулі Advanced Linear / Nonlinear Models / Time Series / Forecasting (рис.2,3).

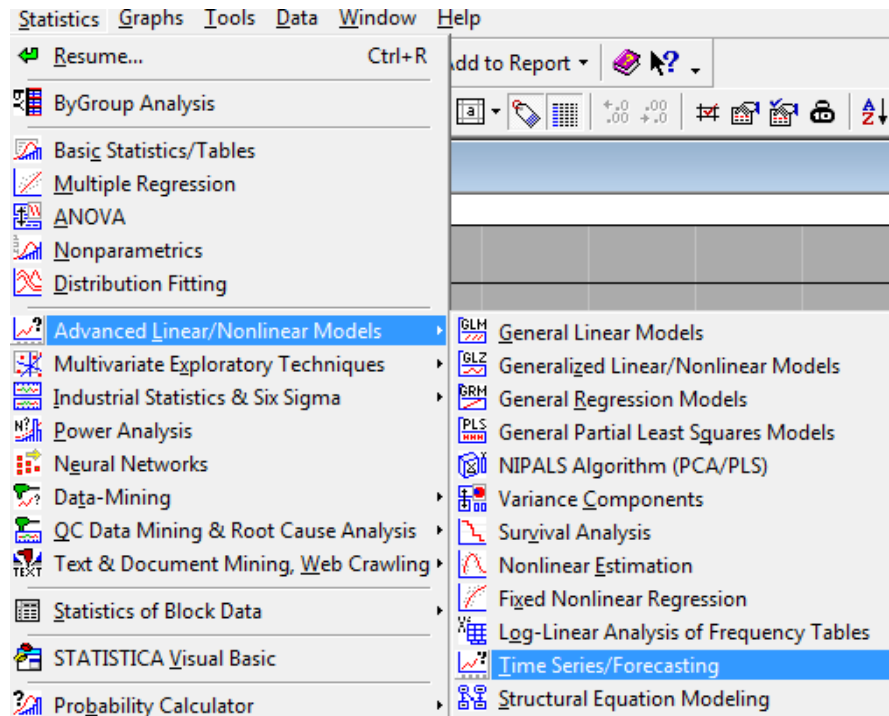


Рисунок 2. Виклик діалогового вікна

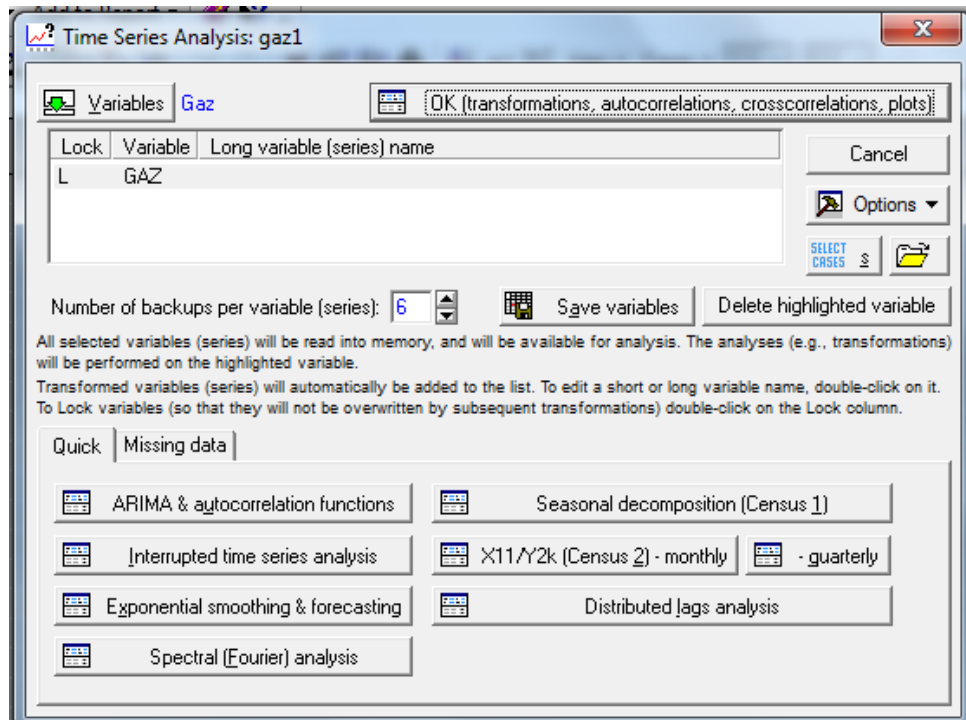


Рисунок 3. Діалогове вікно Time Series Analysis: trend

Перед нами (рис.3) діалогове вікно Time Series Analysis: trend, у верхньому полі якого зберігаються вихідні і перетворені часові ряди. Буква L (в області

Lock) біля назви часового ряду означає, що копія заблокованої змінної не буде видалятися надалі із списку збережених даних. За замовчуванням вихідна змінна позначена заблокованою змінною. Якщо в області збережених змінних є декілька часових рядів, то ряд, з яким будемо працювати. Необхідно виділити.

Кількість копій на один ряд користувач задає самостійно за допомогою перемикача Number of backups per variable (series). За замовчуванням програма робить 6 копій.

Представимо дані графічно за допомогою діалогового вікна Time Series Analysis: trend (рис.4)

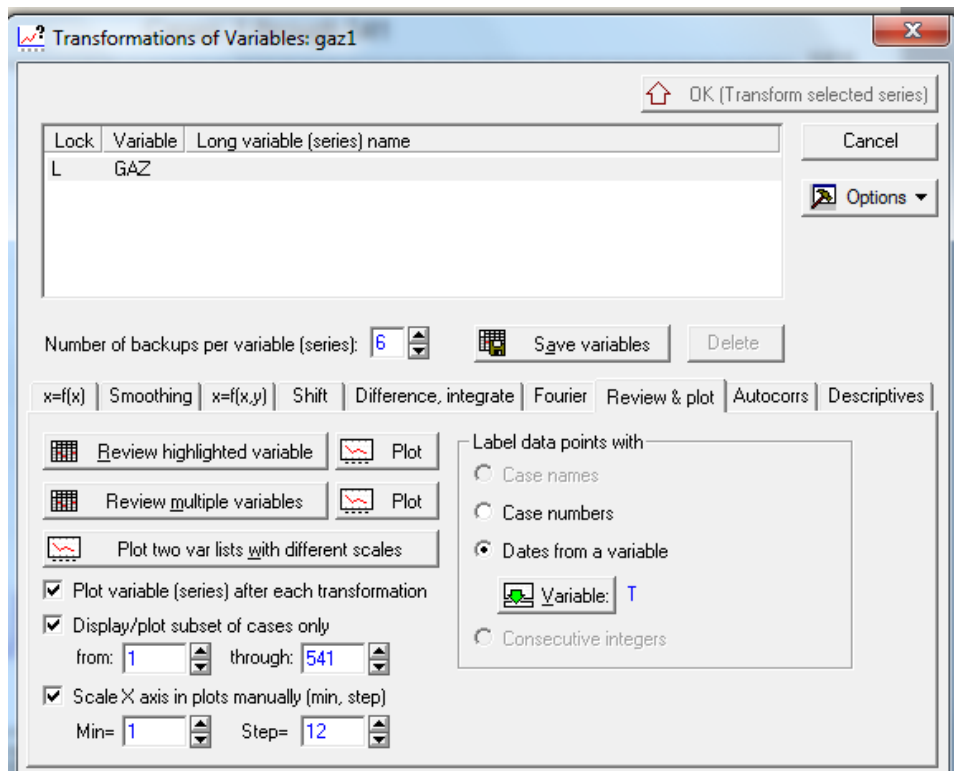


Рисунок 4.

У діалоговому вікні потрібно перейти на закладку Review & plot, де активувати Dates from a variable, відображення дат по осі X вибираємо змінну дат. Щоб задати відображення всіх змінних активуємо Display/plot subset of cases only. Щоб дати не накладались обираємо представлення дат від першої з кроком 12. Після чого натискаємо Plot навпроти Review highlighted variable. У робочій книзі отримуємо графік (рис.5).

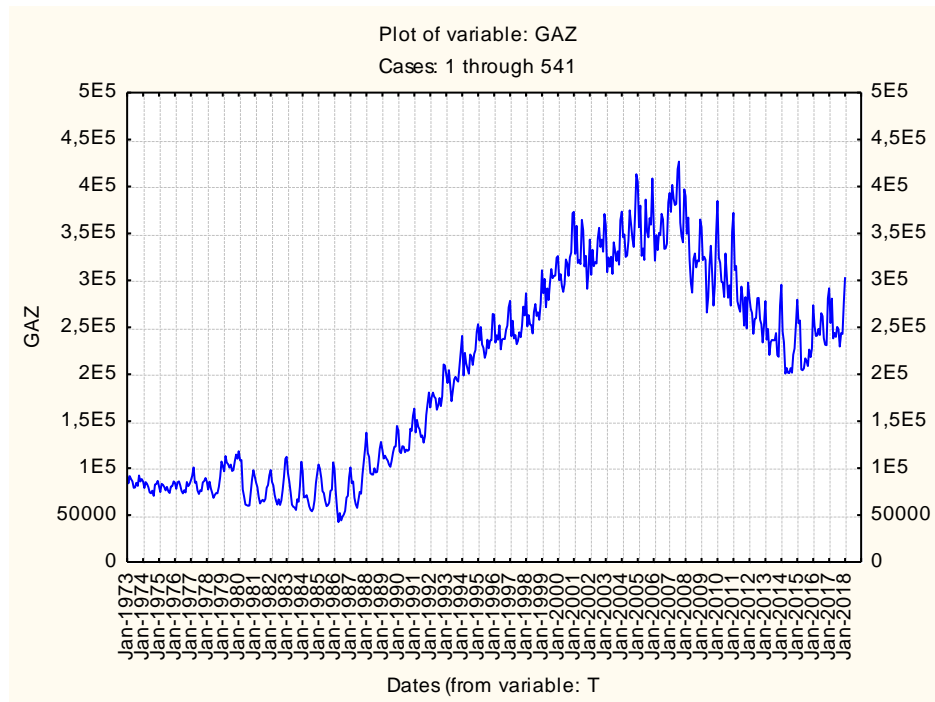


Рисунок 5. Представлення часового ряду

Попередній візуальний аналіз нашого часового ряду показує, що в даних, швидше всього присутній додатній тренд, можлива сезонність і, характер тренду, найімовірніше – мультиплікативний.

Ці припущення варто перевірити аналітично. Зробимо це за допомогою автокореляції. Нагадаймо, що автокореляція дозволяє виявити сезонність і визначити лаги часового ряду. Повертаємося в основне діалогове вікно (рис.4), в якому вибираємо закладку автокореляція (рис.6)

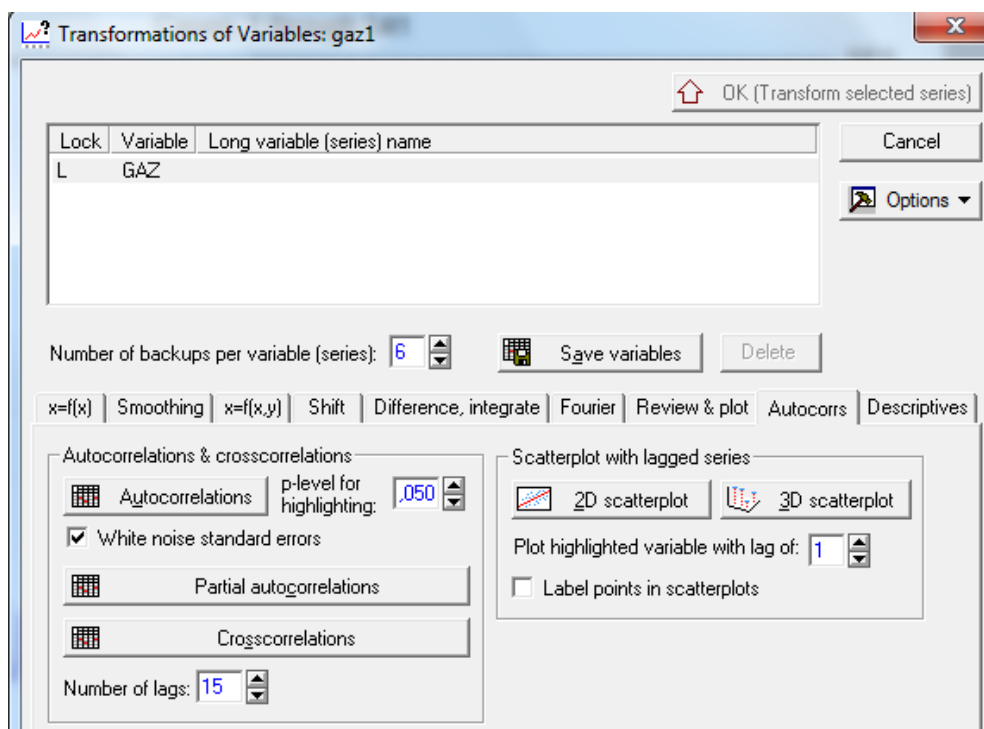


Рисунок 6.

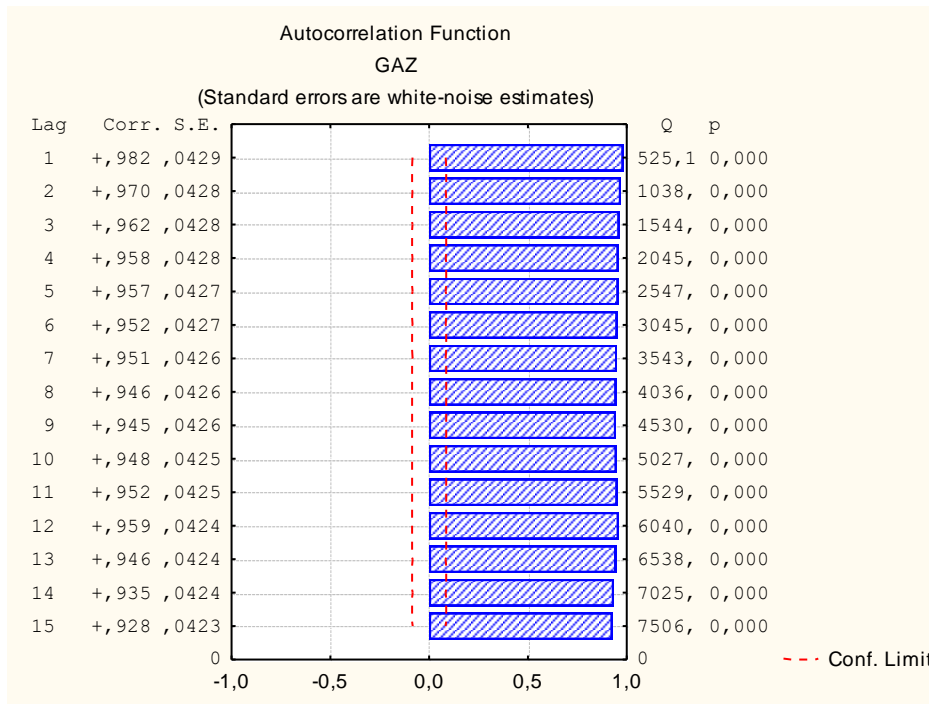


Рисунок 7. Автокореляційна функція

За швидкістю згасання автокореляційної функції можна зробити висновок про характер динаміки. І так ми бачимо, що автокореляційний коефіцієнт високий для перших лагів, далі незначно спадає і досягає свого наступного локального максимуму на 12 лагу. Це і означає, що в наших даних присутній тренд і сезонність із лагом 12. Оскільки відстань між піками зростає (рис.5), то модель може бути мультиплікативною

Нагадаймо, що декомпозицію часового ряду поділяють на такі складові: трендовоциклічна, сезонна і випадкова. Це здійснюють з використанням мультиплікативної моделі часового ряду, яку в загальному вигляді можна представити:

$$Y = T_t \times S_t \times C_t \times I_t$$

Така умовна конструкція дає змогу при прогнозуванні зводити прогнози різних елементів в один кінцевий прогноз.

Для реалізації в пакеті Statistica потрібно вибрати вкладку Seasonal Decomposition у стартовій панелі модуля Advanced Linear / Nonlinear Models / Time Series / Forecasting і задати параметри сезонної декомпозиції (рис. 8, 9).

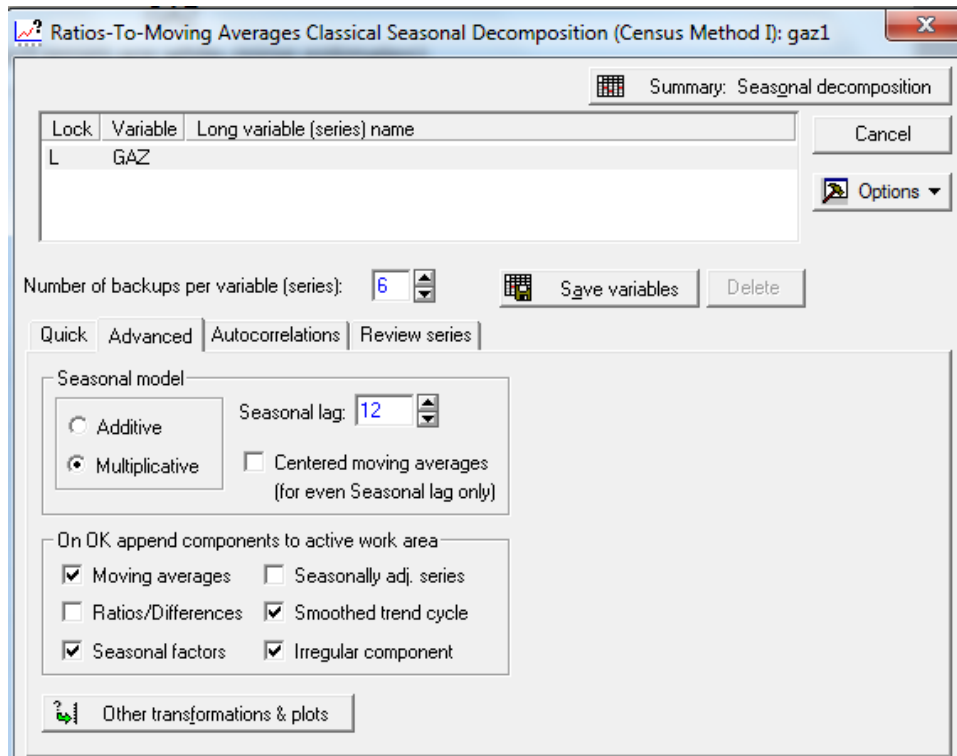


Рисунок 8. Задання параметрів моделі сезонної декомпозиції

Seasonal Decomposition: Multipl. season (12) (gaz1)							
GAZ							
T (Dates)	GAZ	Moving Averages	Ratios	Seasonal Factors	Adjusted Series	Smoothed Trend-c.	Irreg. Compon.
Jan-1973	92694,0			115,2742	80411,8	81397,6	0,987889
Feb-1973	83870,0			102,2188	82049,5	83187,6	0,986319
Mar-1973	91581,0			105,1428	87101,5	86767,5	1,003849
Apr-1973	88407,0			96,1640	91933,5	89088,3	1,031937
May-1973	85844,0			94,1009	91225,5	89186,3	1,022864
Jun-1973	79121,0			90,6407	87290,8	88003,2	0,991905
Jul-1973	79428,0	86075,3	92,2774	95,1799	83450,4	86848,9	0,960869
Aug-1973	84400,0	85502,0	98,7111	94,9691	88871,0	87729,6	1,013011
Sep-1973	81157,0	85089,4	95,3785	92,4977	87739,5	88376,8	0,992789
Oct-1973	92151,0	84556,8	108,9811	98,7870	93282,5	87893,1	1,061318
Nov-1973	85784,0	84137,2	101,9573	101,6214	84415,3	84223,9	1,002273
Dec-1973	88466,0	83630,5	105,7820	113,4033	78010,1	80248,6	0,972105
Jan-1974	85815,0	83224,2	103,1131	115,2742	74444,2	77689,2	0,958231
Feb-1974	78919,0	82709,1	95,4176	102,2188	77205,9	78584,4	0,982459
Mar-1974	85400,0	82010,0	102,8740	105,1428	81022,4	81120,0	0,988904

Рисунок 9. Результат сезонної декомпозиції (фрагмент)

Для подальшого виконання лабораторної роботи необхідно скопіювати результати декомпозиції, а саме трендово-циклічну, сезонну і випадкову складові у вікно з вихідними даними, провести їх специфікацію (рис. 10).

	1 T	2 Gaz	3 Smoothed Trend-c.	4 Seasonal Factors	5 Irreg. Compon.
1	Jan-1973	92694	81397,6	115,2742	0,987889
2	Feb-1973	83870	83187,6	102,2188	0,986319
3	Mar-1973	91581	86767,5	105,1428	1,003849
4	Apr-1973	88407	89088,3	96,1640	1,031937
5	May-1973	85844	89186,3	94,1009	1,022864
6	Jun-1973	79121	88003,2	90,6407	0,991905
7	Jul-1973	79428	86848,9	95,1799	0,960869
8	Aug-1973	84400	87729,6	94,9691	1,013011
9	Sep-1973	81157	88376,8	92,4977	0,992789
10	Oct-1973	92151	87893,1	98,7870	1,061318
11	Nov-1973	85784	84223,9	101,6214	1,002273
12	Dec-1973	88466	80248,6	113,4033	0,972105
13	Jan-1974	85815	77689,2	115,2742	0,958231

Рисунок 10. Результати декомпозиції у вікні вихідних даних (фрагмент)

Далі будемо графік трендово-циклічної складової (рис. 11 - 13)

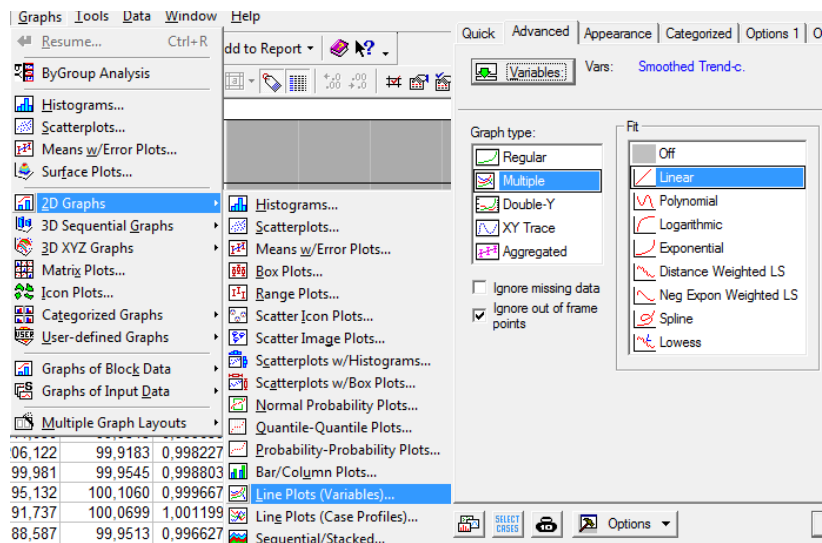


Рисунок 11. Етапи побудови графіка (класично)

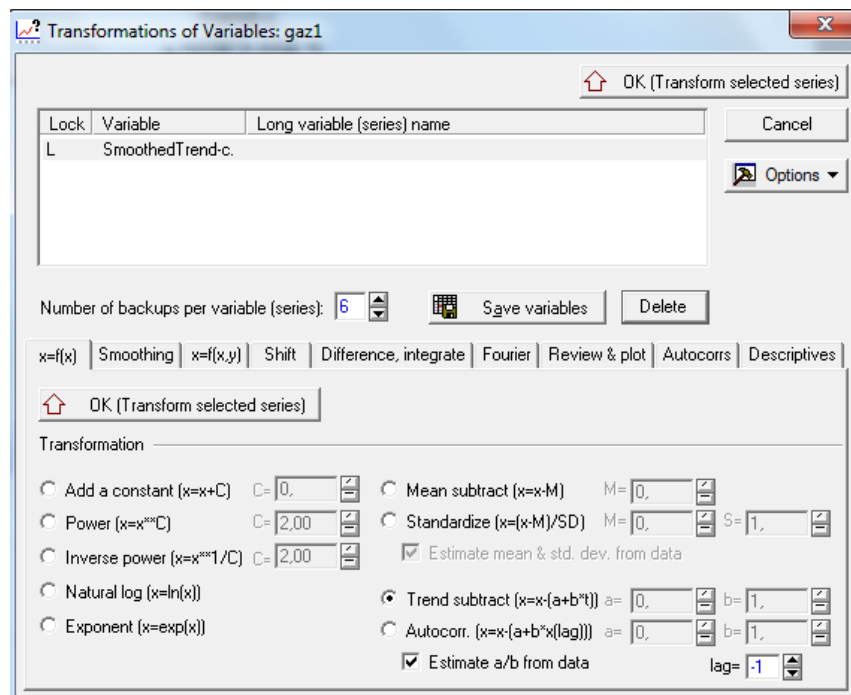


Рисунок 12. Етапи побудови графіка з використанням Time Series /
Forecasting

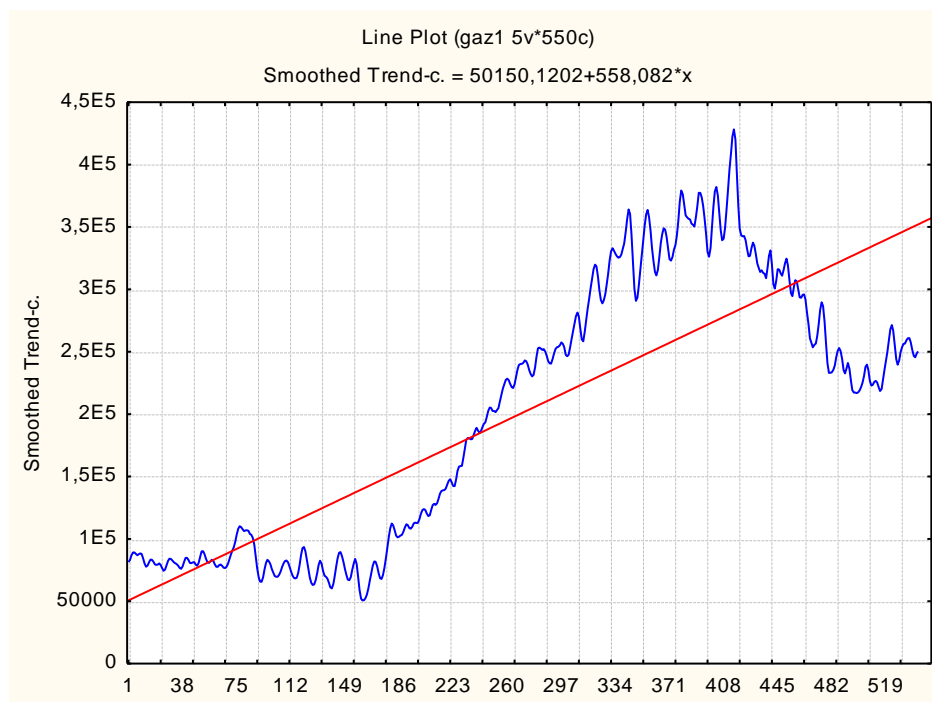


Рисунок 13. **Графік трендово-циклічної складової**

Далі необхідно побудувати графік сезонної складової (рис. 14).

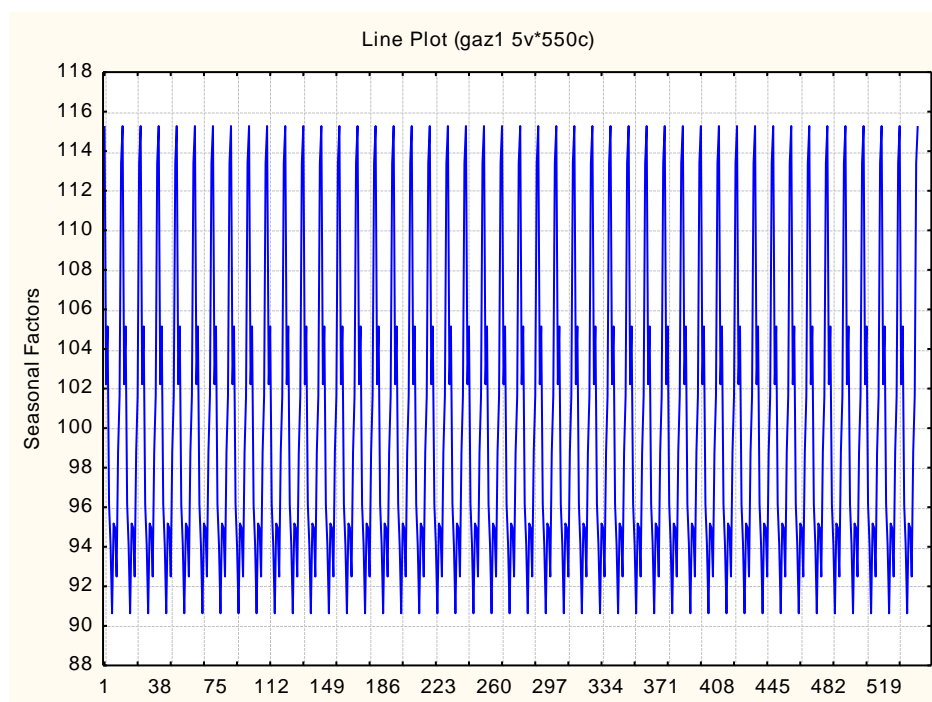


Рисунок 14. **Графік сезонної складової**

Та побудувати графік випадкової складової (рис. 15)

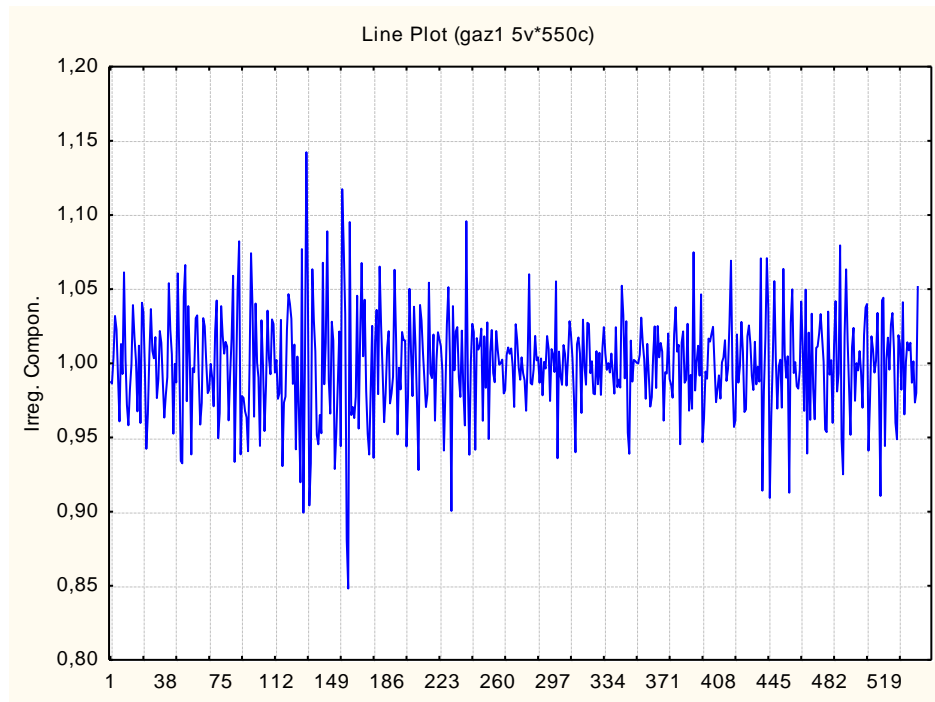


Рисунок 15. Графік випадкової складової

На наступному етапі виділяємо тренд з трендово-циклічної складової. Для цього необхідно додати нову змінну і формулу її розрахунку (рис. 16, 17).

Variable 7

Font: Arial, Size: 10, Bold, Italic, Underline, Superscript, Subscript, Color

Name: Type: OK

Measurement Type: Length: Cancel

Excluded Label Case State MD code: << >>

Display format

- General
- Number
- Date
- Time
- Scientific
- Currency
- Percentage
- Fraction
- Custom

All Specs...
Text Labels...
Values/Stats...
Properties...

Long name (label or formula with): Function guide

Labels: use any text. Formulas: use variable names or v1, v2, ..., v0 is case #.
Examples: (a) = mean(v1:v3, sqrt(v7), AGE) (b) = v1+v2; comment (after:)

Рисунок 16. Вікно введення нової змінної

	1 T	2 Gaz	3 Smoothed Trend-c.	4 Seasonal Factors	5 Irreg. Compon.	6 t	7 Trend
1	Jan-1973	92694	81397,6	115,2742	0,987889	1	151108,202
2	Feb-1973	83870	83187,6	102,2188	0,986319	2	151666,284
3	Mar-1973	91581	86767,5	105,1428	1,003849	3	152224,366
4	Apr-1973	88407	89088,3	96,1640	1,031937	4	152782,448
5	May-1973	85844	89186,3	94,1009	1,022864	5	153340,53
6	Jun-1973	79121	88003,2	90,6407	0,991905	6	153898,612
7	Jul-1973	79428	86848,9	95,1799	0,960869	7	154456,694
8	Aug-1973	84400	87729,6	94,9691	1,013011	8	155014,776
9	Sep-1973	81157	88376,8	92,4977	0,992789	9	155572,858
10	Oct-1973	92151	87893,1	98,7870	1,061318	10	156130,94
11	Nov-1973	85784	84223,9	101,6214	1,002273	11	156689,022
12	Dec-1973	88466	80248,6	113,4033	0,972105	12	157247,104
13	Jan-1974	85815	77689,2	115,2742	0,958231	13	157805,186

Рисунок 17. Результати формування трендової складової (фрагмент)
Необхідно побудувати графік трендової складової (рис. 18).

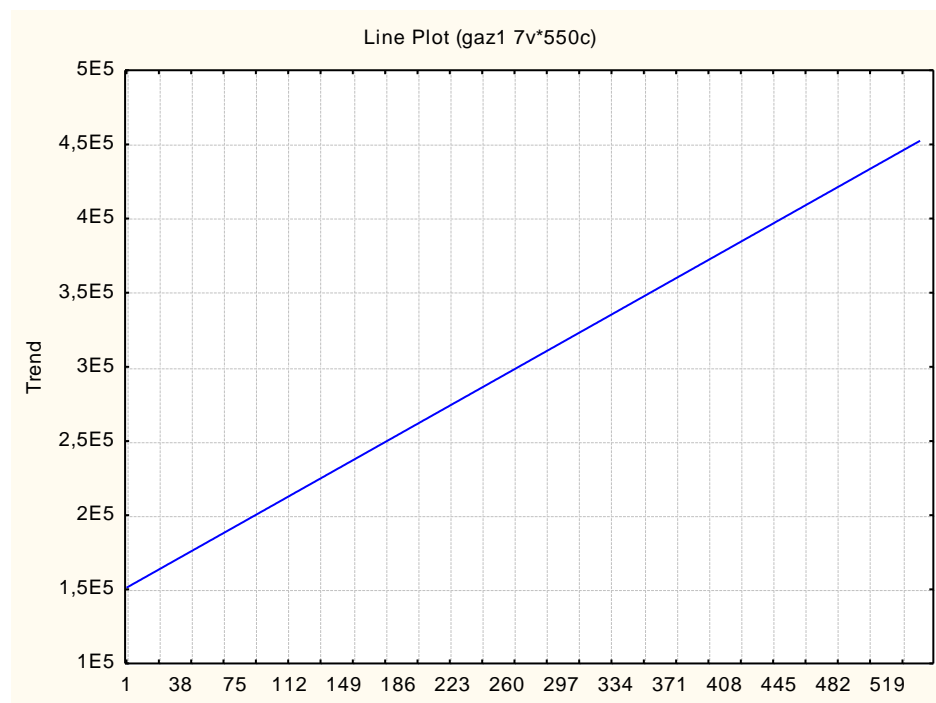


Рисунок 18. Графік тренда

Далі потрібно розрахувати значення циклічної складової, що робимо таким чином (рис. 19 – 21):

$$\text{Cycle} = \text{Smoothed Trend} - C. / \text{Trend.}$$

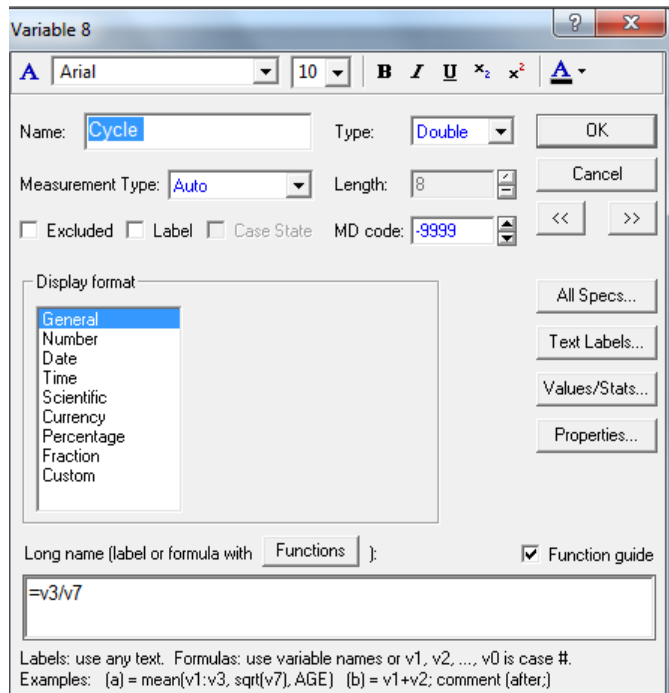


Рисунок 19. Вікно введення нової змінної для циклічної складової

	1 T	2 Gaz	3 Smoothed Trend-c.	4 Seasonal Factors	5 Irreg. Compon.	6 t	7 Trend	8 Cycle
1	Jan-1973	92694	81397,6	115,2742	0,987889	1	151108,202	0,538671
2	Feb-1973	83870	83187,6	102,2188	0,986319	2	151666,284	0,548491
3	Mar-1973	91581	86767,5	105,1428	1,003849	3	152224,366	0,569998
4	Apr-1973	88407	89088,3	96,1640	1,031937	4	152782,448	0,583106

Рисунок 20. Результати розрахунку циклічної складової (фрагмент)

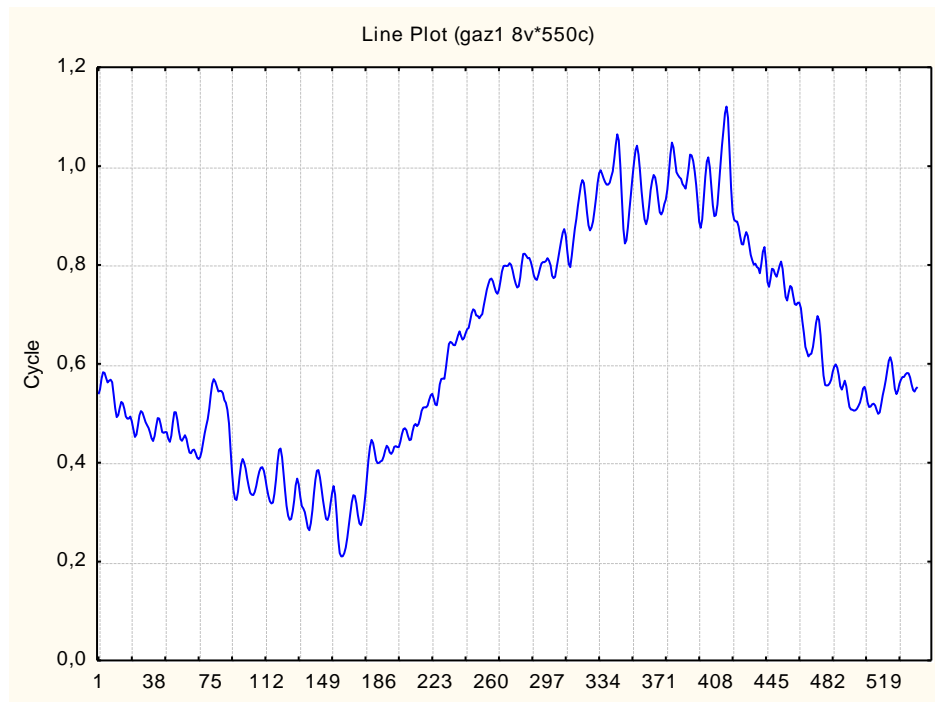


Рисунок 21. Графік циклічної складової

Щоб перейти безпосередньо до прогнозування нашого показника на 12 періодів вперед за допомогою моделі декомпозиції часового ряду, маємо виконати такі дії:

- додати 6 спостережень після останнього з наявних у ряді;
- у стовпці даних t (період часу) вписати відповідні числа, продовжуючи ряд;
- у стовпці Seasonal Factors вписати відповідні значення сезонних складових;
- у стовпці Cycle вписати відповідні значення циклічної складової з урахуванням періоду циклу;
- у стовпці Trend задати перерахунок даних;
- додати нову змінну Obs_predict.

Тоді розрахувати прогнозні значення показника на 12 кроків вперед можна шляхом вказівки моделі виду (рис. 22):

$$\text{GAZpredict} = \text{Trend} \times \text{Cycle} \times \text{Seasonal Factors} / 100.$$

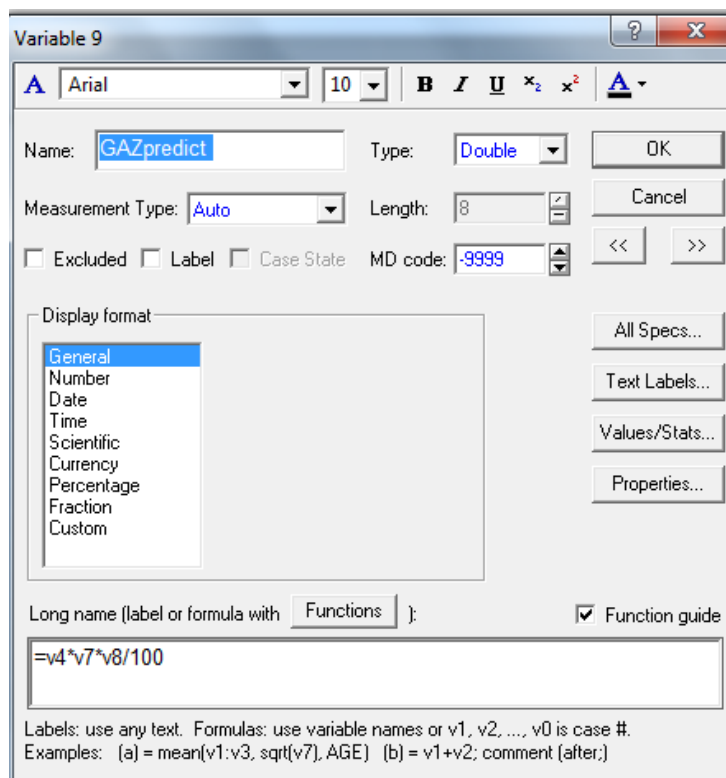


Рисунок 22. Вікно введення нової змінної

Прогнозні значення імпорту природного газу наведено на рис. 23, а його графік – на рис. 24

	1 T	2 Gaz	3 Smoothed Trend-c.	4 Seasonal Factors	5 Irreg. Compon.	6 t	7 Trend	8 Cycle	9 iAZpredic
1	Jan-1973	92694	81397,6	115,2742	0,987889	1	151108,202	0,538671	93830,42
2	Feb-1973	83870	83187,6	102,2188	0,986319	2	151666,284	0,548491	85033,37
3	Mar-1973	91581	86767,5	105,1428	1,003849	3	152224,366	0,569998	91229,84
4	Apr-1973	88407	89088,3	96,1640	1,031937	4	152782,448	0,583106	85670,93
539	Nov-2017	242922	245535,6	101,6214	0,973570	539	451356,318	0,543995	249516,8
540	Dec-2017	276341	248664,0	113,4033	0,979956	540	451914,4	0,550246	281993,2
541	Jan-2018			115,2742		541	452472,482	0,559968	292070,6
542				102,2188		542	453030,564	0,568162	263105,9
543				105,1428		543	453588,646	0,573512	273517,2
544				96,1640		544	454146,728	0,573506	250464,9
545				94,1009		545	454704,81	0,578618	247579,8
546				90,6407		546	455262,892	0,581545	239976,6
547				95,1799		547	455820,974	0,580953	252046,5
548				94,9691		548	456379,056	0,572549	248153,6
549				92,4977		549	456937,138	0,55835	235990,3
550				98,7870		550	457495,22	0,547159	247286,1
551				101,6214		551	458053,302	0,543995	253219
552				113,4033		552	458611,384	0,550246	286172,1

Рисунок 23. Прогнозне значення показника (фрагмент розрахункової таблиці)

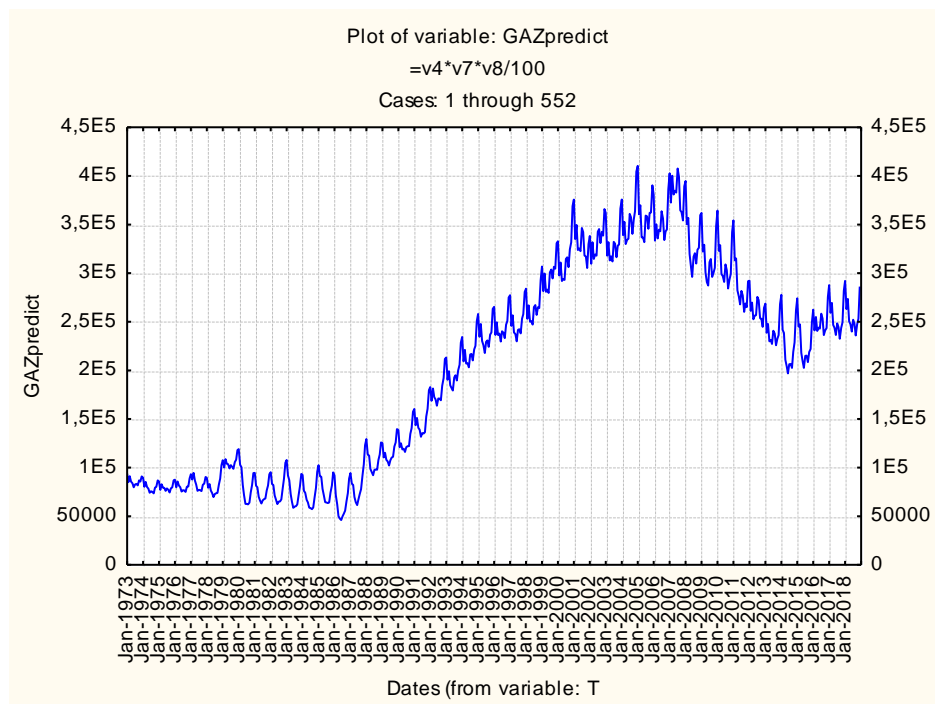


Рисунок 24. Графік прогнозних значень показника

На рис. 25 наведено гістограму розподілу похибок. Те, що зазначений розподіл близький до нормального закону, є підтвердженням адекватності моделі і точності прогнозу.

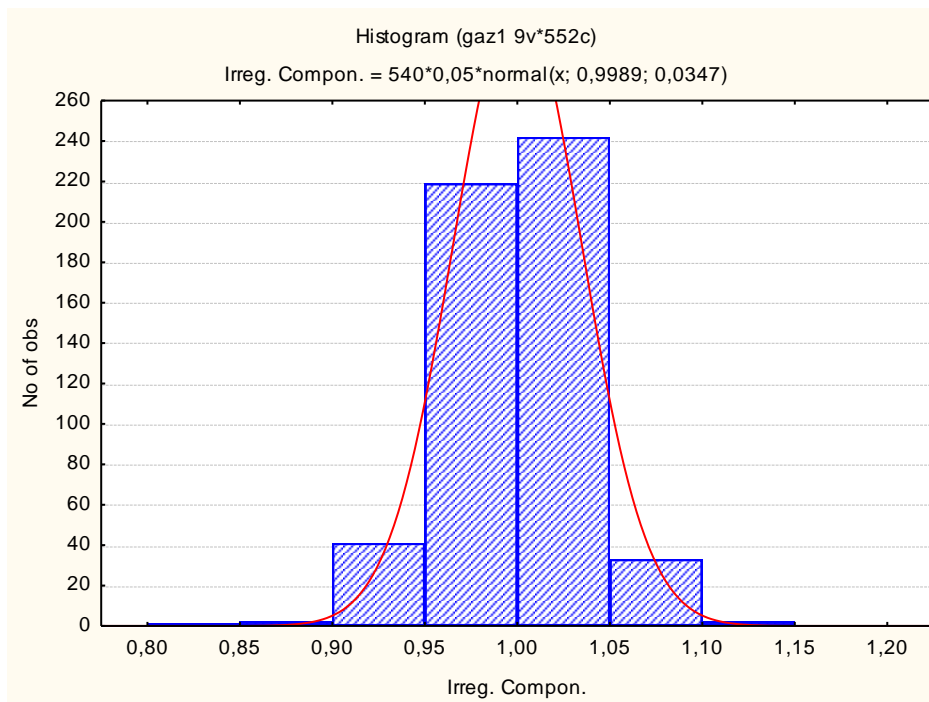


Рисунок 25. Гістограма розподілу помилок

Слід розрахувати середню відносну відсоткову похибку за формулою (рис. 26):

$$\varepsilon = \frac{1}{n} \sum \frac{(y_i - y_{ti})}{y_i} 100 \% \quad (2)$$

529	Jan-2017	291656	249620,2	115,2742	1,013582	529	445775,498	0,559968	287747,7	0,013
530	Feb-2017	254597	253589,8	102,2188	0,982179	530	446333,58	0,568162	259216,5	0,01814
531	Mar-2017	280567	256297,8	105,1428	1,041147	531	446891,662	0,573512	269478,8	0,03952
532	Apr-2017	238279	256615,1	96,1640	0,965586	532	447449,744	0,573506	246771,5	0,03564
533	May-2017	244126	259225,4	94,1009	1,000789	533	448007,826	0,578618	243933,4	0,00078
534	Jun-2017	239805	260861,3	90,6407	1,014204	534	448565,908	0,581545	236446,5	0,01400
535	Jul-2017	250508	260920,0	95,1799	1,008716	535	449123,99	0,580953	248343,4	0,00864
536	Aug-2017	247888	257464,9	94,9691	1,013806	536	449682,072	0,572549	244512,2	0,01361
537	Sep-2017	229482	251391,7	92,4977	0,986885	537	450240,154	0,55835	232531,6	0,01328
538	Oct-2017	243985	246658,2	98,7870	1,001308	538	450798,236	0,547159	243666,2	0,00130
539	Nov-2017	242922	245535,6	101,6214	0,973570	539	451356,318	0,543995	249516,8	0,02714
540	Dec-2017	276341	248664,0	113,4033	0,979956	540	451914,4	0,550246	281993,2	0,02045
541	Jan-2018			115,2742		541	452472,482	0,559968	292070,6	
542	Feb-2018			102,2188		542	453030,564	0,568162	263105,9	
543	Mar-2018			105,1428		543	453588,646	0,573512	273517,2	
544	Apr-2018			96,1640		544	454146,728	0,573506	250464,9	
545	May-2018			94,1009		545	454704,81	0,578618	247579,8	
546	Jun-2018			90,6407		546	455262,892	0,581545	239976,6	
547	Jul-2018			95,1799		547	455820,974	0,580953	252046,5	
548	Aug-2018			94,9691		548	456379,056	0,572549	248153,6	
549	Sep-2018			92,4977		549	456937,138	0,55835	235990,3	
550	Oct-2018			98,7870		550	457495,22	0,547159	247286,1	
551	Nov-2018			101,6214		551	458053,302	0,543995	253219	
552	Dec-2018			113,4033		552	458611,384	0,550246	286172,1	
MEAN case 529-540										0.01716

Рисунок 26. Розрахунок середньої відносної відсоткової похибки (фрагмент розрахункової таблиці)

Середня відносна відсоткова похибка рівна 1,7163 %, що підтверджує точність прогнозу за даною мультиплікативною моделлю часового ряду.