



Міністерство освіти і науки України

**Національний технічний
університет**

"Дніпровська політехніка"

І.М. Пістунов

ЗБІРНИК ІНДИВІДУАЛЬНИХ ЗАВДАНЬ

з дисципліни

«DATAMANING»

Дніпро

НТУ «ДП»

2025

УДК 004.738.5:338.46(075)
ПЗ4

Затверджено вченою радою університету як навчальний посібник по дисципліні «Економіко-математичне моделювання» для студентів очної та заочної форм навчання зі спеціальності 051 Економіка (Протокол № від р).

Рецензенти:

А.В. Бардась, докт. екон. наук, проф., декан факультету менеджменту Національного гірничого університету;
Н.К. Васильєва, докт. екон. наук, проф., -завідувач кафедри інформаційних систем і технологій Дніпровського агро-економічного університету.

Пістунов І.М., Демиденко М.А.

ПЗ4 Збірник індивідуальних завдань для дисциплін «Datamining» [Електронний ресурс]: Методичні вказівки. Дніпро: Державний НТУ «ДП», 2025. 25 с. Режим доступу: http://pistunovi.inf.ua/Datamining_Ind_task.pdf (дата звернення: 26.12.2024). – Назва з екрана.

В посібнику розглядається основні методи дейтамайнінгу, які включають в себе такі дисципліни як інформатика та комп'ютерна техніка, економетрика, теорія ймовірності, економічна кібернетика, кластерний аналіз, нейронні сітки, прогнозування часових рядів.

Наведено приклади розрахунків із застосування електронних таблиць Excel.

Призначено для студентів всіх спеціальностей.

Збірник скомпоновано для практичних чи лабораторних занять із застосуванням комп'ютерної техніки.

Призначений для студентів вищих учбових закладів і може бути корисним для фінансистів, економістів, плановиків, менеджерів та маркетологів.

Посібник базується на літературних джерелах вітчизняних, зарубіжних авторів, ресурсах Інтернету та на досвіді викладання дисципліни «Смарт-технології для економії» в Державному НТУ «ДП».

© І.М. Пістунов, М.А. Демиденко, 2025
© Державний НТУ « ДП », 2025

ЗМІСТ

1. ВКАЗІВКИ ДО РОБОТИ З EXCEL	4
2. ПРИКІНЦЕВІ ПОЛОЖЕННЯ1	3
3. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ №1. Аналіз кошику продажів...	
4. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ №2. «Дослідження купівельної спроможності українців з 1950 до 2021рр».....	
5. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ №3. Синтез квазілінійних моделей	
6. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ №4. Кластеризація об'єктів.....	
7. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ №5. Прогнозування і методи очищення даних.....	

1. ВКАЗІВКИ ДО РОБОТИ З EXCEL

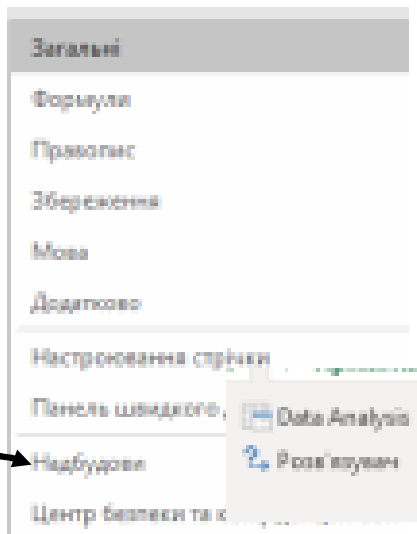
1. Ті завдання, які студенти виконують із застосуванням табличного процесора Excel з версії Microsoft Office не пізніше 2007 року.

1. Числові значення кожного завдання обираються з таблиць, вміщених для кожного завдання окремо.

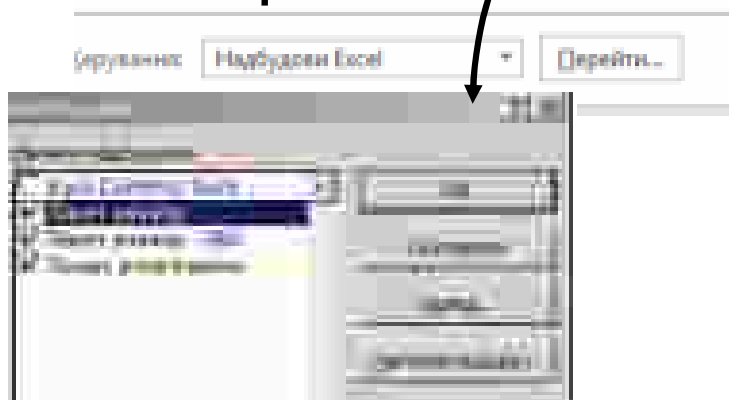
2. На сторінці процесора спочатку вставляється текст задачі, потім, значення букв, потім формула, за якою вирішується задача, далі рішення і текстові висновки.

3. Перед початком роботи потрібно налаштувати Microsoft Excel. Доступ до надбудов здійснюється через правому кутку можна

Після інсталяції Microsoft програма Excel, необхідно програми для розрахунків оптимізації. Для цього «Файл-Параметри». В меню,



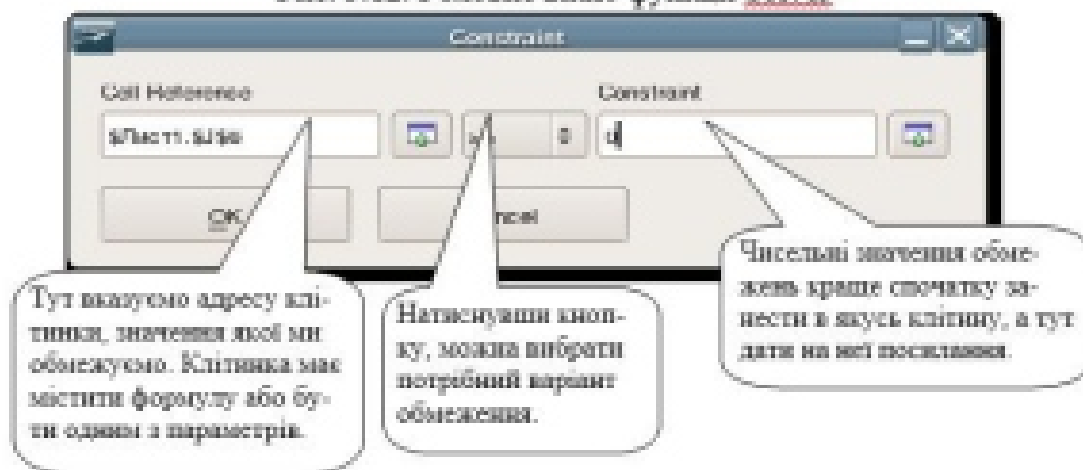
головне меню «Дані» і у побачити ці надбудови . Office, куди входить і провести налаштування цієї коефіцієнтів моделей та потрібно вибрати пункт що з'явиться, вибрати пункт «Надбудови».



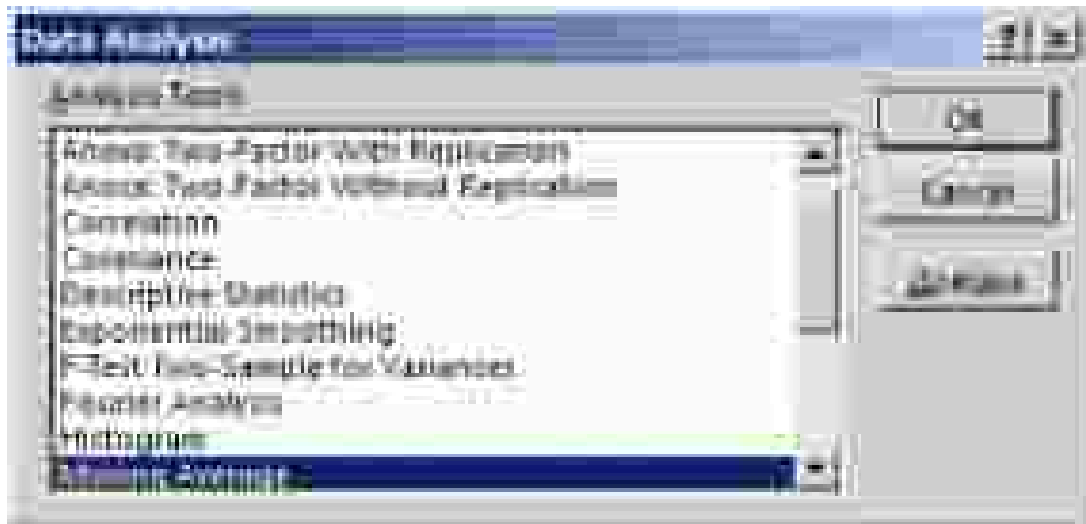
Далі клацнути на кнопку «перейти» та відмітити вказані на рисунку позиції і натиснути «ОК» у цьому і наступних вікнах. Тепер в головному меню програми за пунктом «Дані» у правому кутку меню з'являться пункти «Data Analysis» та «Розв'язувач».



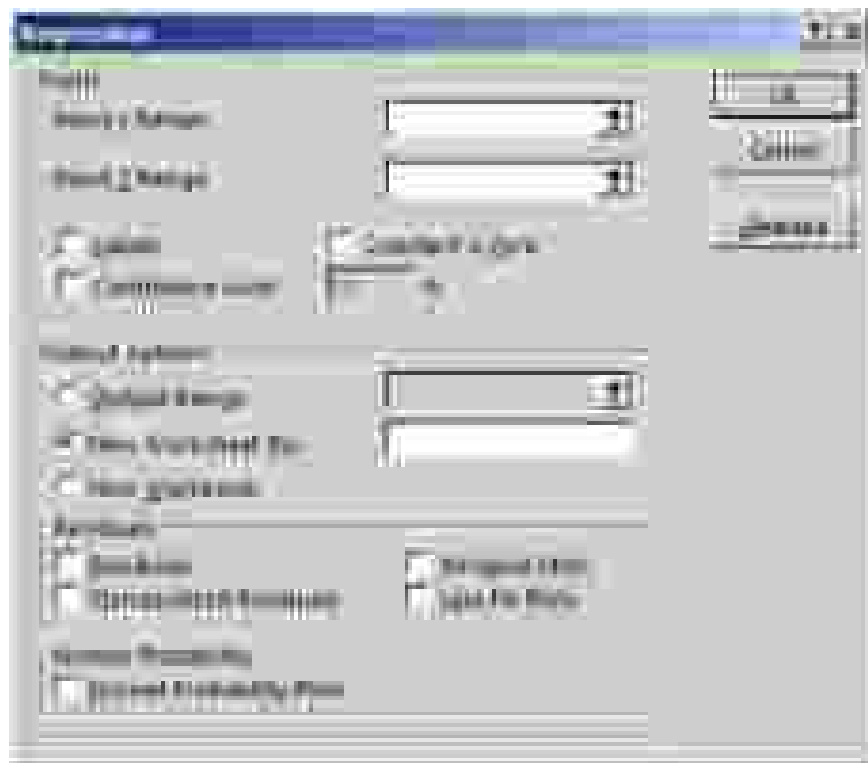
Рис. 3.12. Головне вікно функції Solver



4. При використанні функції «Аналіз даних» ви обираєте зі списку потрібний вид аналізу



Наприклад, якщо обираєте пункт «Регресія», то побачите наступне вікно



Порядок користування різними підпрограмами вам буде роз'яснено на практичних заняттях.

5. Завдання треба здавати в електронному вигляді на будь яких носіях у конвертах, які потрібно підписувати таким чином. Допускається здавати всі завдання на одному носії.

Задачу спочатку треба розв'язати в загальному вигляді з представлення формули рішення, в яку потім підставлені конкретні числові дані для свого варіанта. В деяких завданнях числові значення потрібно визначити за простою формулою. Наприклад, якщо в таблиці навпроти позначення C стоїть число 17, а числове значення в умові задачі подано як $0,01 \cdot C$, то це означає, що потрібно брати число $0,01 \cdot 17 = 0,17$.

Кожну тему супроводжують приклади вирішення із застосування таких прикладних пакетів як Open Office Calc, Microsoft Office Excel, Macsima та MathCad. При цьому припускається, що студенти вже знайомі з порядком використання як електронних таблиць так і математичних процесорів.

Наприклад, якщо потрібно провести розрахунки за формулою $A = \frac{B - C}{D}$

для наступних числових значень параметрів $B = 10$, $C = 5$, $D = 8$, то в підрозділі буде наведено малюнок, в якому видно фрагмент вікна електронної таблиці, де колонку A займають тестові визначення невідомих у формулі, колонку B – їх числові значення. Вікно f_x містить саму формулу розрахунку, де вказано адреси клітинок, які містять числові дані.

	A	B	C	D
1	B=	10		
2	C=	5		
3	D=	8		
4	A=	0.625		

Якщо будуть застосовані функції електронних таблиць, то буде показано їх вікно з уведеними туди параметрами

2. ПРИКІНЦЕВІ ПОЛОЖЕННЯ

Метою даного курсу є практичне засвоєння студентами методів та прийомів, які дозволяють економістам знайти найкращі рішення у практичній діяльності. Для цього використовуються Microsoft Excel.

Звіт з кожного індивідуального завдання може подаватися тільки в електронному вигляді у форматі *.XLSX. Ім'я файлу – це номер групи та прізвище й ініціали. Наприклад, 051-24-1_Петренко В.О.xlsx. Файли розміщуються у «хмарному» середовищі, яке вказує викладач.

Звіт має містити:

1. Опис завдання.
2. Початкові значення
3. Результат розрахунку
4. Зображення активних вікон спеціальних програм.
5. Висновки.

Останній пункт є найважливішим, оскільки головним у цьому курсі є не вміння провести розрахунки, але розуміння, для чого ці розрахунки проводяться, що означають з економічної точки зору отримані результати, де можна застосувати кожен конкретний метод і яку вигоду він принесе в реальній економічній діяльності. Відповіді на всі ці питання і має містити висновок.

Якщо у завданні відсутні ціни та інші показники, студенти мають знайти їх пошуком в Інтернеті.

3. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ№1.

Аналіз кошику продажів.

Критерії оцінювання: це завдання оцінюється у 5 балів за національною шкалою. За кожну помилку знімається 0,1 бали. Потім оцінка перераховується за 100-бальною системою згідно існуючого положення.

Мета роботи: Набути навичок в сортуванні та фільтрації даних та аналіз кошику продажів комерційного підприємства.

- Завдання:**
1. Скачати файл лабораторна робота _1.xlsx.
 2. У клітинку A2 вставити формулу = RANDBETWEEN(0;1).
 3. Розтягнути цю формулу на всю таблицю і натиснути Ctrl + C.
 4. По головному меню обрати пункт «Вставити- Значення».
 5. Виконати аналіз та оформлення результатів користуючись прикладом розрахунку з файлу Лр_1_приклад.xlsx

4. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ № 2.

«Дослідження купівельної спроможності українців з 1950 до 2021рр»

Критерії оцінювання: це завдання оцінюється у 5 балів за національною шкалою. За кожну помилку знімається 0,1 бали. Потім оцінка перераховується за 100-бальною системою згідно існуючого положення.

Мета роботи: Набути навичок у визначенні справжнього рівня купівельної спроможності із використанням табличного процесору Excel. Дослідити як змінювалась спроможність населення України задовольняти свої потреби в споживчих товарах на середню заробітну платню.

Завдання: 1. Для окремих років відомі знайти в інтернеті середні зарплати прикладом такої інформації може бути такий [сайт](#).

https://uk.m.wikipedia.org/wiki/%D0%A1%D0%B5%D1%80%D0%B5%D0%B4%D0%BD%D1%8F_%D0%B7%D0%B0%D1%80%D0%BE%D0%B1%D1%96%D1%82%D0%BD%D0%B0_%D0%BF%D0%BB%D0%B0%D1%82%D0%B0_%D0%B2_%D0%A3%D0%BA%D1%80%D0%B0%D1%97%D0%BD%D1%96

Або такий [сайт https://www.pfu.gov.ua/2130793-pokaznyk-serednoyi-zarobitnoyi-platy-za-2021-rik/](https://www.pfu.gov.ua/2130793-pokaznyk-serednoyi-zarobitnoyi-platy-za-2021-rik/) Наприклад ціни у 1950р, за [посиланням https://www.russian-money.ru/prices/?year=1950](https://www.russian-money.ru/prices/?year=1950), додатково по інших роках <https://argumentua.com/reportazh/deshevizna-v-sssr-eto-obyknovennyi-mif-tseny-na-tovary-togda-v-pereschete-na-segodnya>.

2. Знайти в Інтернеті ціни на товари у відповідному році - **Н**.
3. Знайти в Інтернеті склад споживчого кошика на 1 місяць.
4. Ділимо ціну одиниці окремого товару **Р**, який входить до споживчого кошика на 1 місяць, на середню заробітну плату.

5. Знаходимо частку Δ зарплати яку треба витратити на один ти товару зі споживчого кошика у P в році N .
6. Обчислюємо частки Δ по наступних періодах дослідження (роках). Такі обчислення виконуємо не по одному товару, а по групі однакових продовольчих і побутових товарів, дивись табл. 2. Для зручності, всі ціни варто привести до поточного курсу долара.
7. Отримавши результати будуємо таблиці і графіки.
8. Формулюємо висновки в яких показуємо як змінилась купівельна спроможність. чи подешевшали товари відносно заробітної плати, чи подорожчали. Які товари, або групи товарів подешевшали чи подорожчали. Чому такі зміни мали місце, на ваш погляд. Відповідаємо на питання чи зараз стало краще чи гірше відносно досліджуваного періоду.
9. В якості досліджуваних періодів взяти: 1950-1959; 1960-1969, 1970-1979, 1980-1989, 1990-1999, 2000-2009, 2010-2019, 2020-2023.

Таблиця 2

Перелік груп товарів

Остання цифра номера залікової книжки	Продукти
0	Хлібобулочні вироби, макарони
1	М'ясо
2	Яйця
3	Риба
4	Сири
5	Овочі, фрукти
6	Ліки
7	Пальне
8	Одяг
9	Розваги

Вихідні данні : Результати пошуку в Інтернеті

№	Товар	Ціна товару 1950-1959	Ціна товару 1960-69	Ціна товару 1960-69
1				
2				

Програмне забезпечення для обчислення : MS Excel

Звіт по завданню:

1. Постановка задачі дослідження;
2. Пошук вихідних даних. Інформація подається у вигляді таблиць MS Excel. Таблиці супроводжуються посиланням на джерело в Інтернеті;
3. Розрахунки по дослідженню.
4. Графіки, таблиці, малюнки, які пояснюють дослідження;
5. Висновки дослідження.
6. Звіт оформлюється у вигляді документа MS Word, розрахунки, що виконані подаються у документі MS Excel.

Додатково: Студенти можуть доповнювати роботу своїми дослідженнями, які вони вважають за доцільне.

5. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ № 3.

Синтез квазілінійних моделей

Критерії оцінювання: це завдання оцінюється у 5 балів за національною шкалою. За кожну помилку знімається 0,1 бали. Потім оцінка перераховується за 100-бальною системою згідно існуючого положення.

Мета роботи: Набути навичок у побудові лінійних та нелінійних моделей методом найменших квадратів із використанням табличного процесору Excel.

- Завдання:**
1. Із табл.2 обрати тип змінної згідно останньої цифри номера залікової книжки, яка буде слугувати ендогенною змінною (Y), а всі інші типи продуктів будуть екзогенними змінними (X)
 2. Вивести дані на аркуш Microsoft Excel не по періодам а по рокам.
 3. Розрахувати кореляційну матрицю зв'язків факторів
 - 4.. Застосувати методикку синтезу лінійних статистичних моделей через додаток Microsoft Excel – Regression, який знаходиться через головне меню ДАНІ – Data Analysis.
 5. Побудувати нелінійну модель, створивши нелінійні ефекти у початковій матриці даних.
 6. Побудувати авторегресійну модель.
 6. Зробити висновки.

6. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ № 4.

Кластеризація об'єктів

Критерії оцінювання: це завдання оцінюється у 5 балів за національною шкалою. За кожну помилку знімається 0,1 бали. Потім оцінка перераховується за 100-бальною системою згідно існуючого положення.

Мета роботи: Вивчити методи розрахунків відстаней між об'єктами за допомогою різних метрики Евкліда та Джеффріса-Матусіти та прийоми по їх автоматичній кластеризації. із використанням табличного процесору Excel.

- Завдання:**
1. Із табл.3 обрати дані згідно останню цифру номеру залікової книжки N_3 .
 2. розрахувати дві матриці відстаней: за метриками Евкліда, та Джеффріса-Матусіти;
 2. Ввести дані на аркуш Microsoft Excel.
 3. Застосувавши функцію Solve електронних таблиць Excel з пакету Office 365, потрібно вирішити оптимальні задачі включення до кластерів для матриць відстаней, розрахованих за цими метриками.
 4. Порівняти результати кластеризації і зробити висновки.

Таблиця 3

Варіанти завдань

№ п/п	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
0	64	0,29	340	2710	304,92	8368,7	4,98953	4,43939	19602
	31	0,73	370	2770	279,51	8483,3	7,67620	3,10606	7194
	39	0,69	470	2780	415,03	4098,3	5,47801	1,90909	9504
	44	0,54	290	1280	145,68	4728,9	4,74529	1,86363	11814
	87	0,64	340	1460	272,73	5674,6	5,61758	1,84848	15378
	74	0,76	360	1220	448,91	2866	8,72295	1,27272	5148
	54	0,54	710	2480	149,07	8110,7	5,1291	4,10606	18084
	74	0,64	770	1790	220,22	7652,2	6,28053	2,80303	13992
	78	0,85	300	1690	150,76	2780,0	9,35101	4,75757	17622
	87	0,39	560	2770	501,42	8311,4	6,07117	3,96969	19074
1	25	0,82	430	2590	164,31	5416,7	4,50104	4,46969	20262
	67	0,29	270	2870	492,95	3525,1	10,8513	2,98484	7326
	62	0,52	860	1920	528,52	4069,7	9,24633	3,19697	13530
	53	0,54	790	2770	133,82	3926,4	4,67550	4,65151	13068
	42	0,42	610	1100	232,07	5474,0	4,29169	1,72727	6270
	64	0,29	700	2860	267,65	7107,6	9,69993	2,45454	14916
	30	0,57	550	2150	531,91	6305,2	8,72295	2,18181	8646
	41	0,88	800	1840	499,73	6276,5	6,28053	2,10606	12342
	68	0,37	740	1220	381,15	6047,2	5,82693	4,50000	20592
	49	0,58	490	2430	474,32	4384,9	4,57083	4,22727	8976
2	76	0,49	750	3100	247,32	5560,0	8,65317	3,62121	11616
	35	0,49	380	1850	152,46	5760,6	4,04745	1,62121	5214
	43	0,38	290	840	138,90	4958,1	10,5722	1,34848	14916
	52	0,57	860	2030	218,52	8769,9	10,9909	3,28787	10164
	72	0,26	620	1450	203,28	8827,2	9,76971	2,43939	6270
	73	0,28	460	1360	470,93	7824,1	6,62944	2,62121	8844
	88	0,37	750	2230	433,66	8712,6	10,8513	3,96969	9438
	42	0,59	770	1520	252,40	7795,5	10,6420	3,92424	5412
	69	0,72	740	990	320,16	3983,7	3,66364	4,45454	6204
	75	0,31	380	2830	135,52	4470,9	8,51360	4,33333	18150
3	33	0,26	390	2420	526,83	2636,7	3,24494	4,09090	19338
	41	0,29	760	1170	164,31	5846,6	8,09490	4,40909	20328
	69	0,75	640	1890	531,91	5216,1	10,1535	1,68181	11352
	66	0,54	860	2880	177,87	5846,6	3,21004	4,07575	13398
	38	0,52	360	2730	238,85	5531,3	3,66364	4,65151	19536
	57	0,5	680	1230	476,01	7709,5	8,44382	3,06060	8316
	58	0,59	410	2620	367,59	2264,1	2,82623	2,16666	7326
	33	0,69	310	1280	513,28	8110,7	5,23377	3,72727	18216
	60	0,89	800	1630	215,13	6878,4	8,79274	2,31818	14058
	32	0,76	500	2370	216,83	2350,1	6,14096	1,40909	17028
4	59	0,46	460	2300	442,13	4528,2	6,38520	3,83333	13464
	68	0,5	470	1560	282,89	3238,5	3,83810	3,09090	8316

№ п/п	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	34	0,34	330	2070	430,27	7050,3	3,03559	2,39393	17754
	38	0,76	400	850	509,89	6706,4	8,19958	3,10606	19866
	56	0,76	890	1140	164,31	5674,6	5,58269	3,54545	5610
	84	0,43	870	1860	203,28	6591,8	4,08234	3,10606	13398
	86	0,65	540	820	166,01	2665,3	5,16399	4,09090	16500
	87	0,6	310	1740	481,09	7594,9	9,66503	3,33333	15840
	60	0,89	250	1230	370,98	7852,8	2,68667	4,43939	19602
	28	0,58	870	1670	210,05	4442,3	6,87369	4,03030	7062
5	86	0,41	510	2140	218,52	8941,9	10,2233	3,31818	17094
	81	0,33	400	1270	226,99	5359,4	5,26866	4,45454	18084
	58	0,74	630	2670	440,44	8769,9	4,67550	2,15151	6336
	40	0,57	490	810	159,23	5703,3	3,38450	2,92424	11352
	55	0,67	460	3060	181,25	4155,7	4,53593	1,98484	9306
	53	0,42	390	920	404,86	8110,7	5,09420	4,07575	9636
	44	0,78	400	2730	272,73	4155,7	4,81507	2,30303	11088
	59	0,79	310	980	143,99	4384,9	8,40893	1,59090	15246
6	29	0,45	890	2230	164,31	4069,7	10,9211	1,60606	16962
	88	0,7	330	1890	274,42	2722,7	5,02442	4,18181	7062
	25	0,42	580	1810	399,78	5101,4	8,12979	4,34848	20790
	69	0,7	800	3040	362,51	5646,0	2,79134	1,54545	10164
	74	0,46	570	2950	459,07	8082,1	3,52407	1,62121	17490
	72	0,52	360	1810	499,73	7766,8	3,83810	3,92424	7062
	86	0,68	790	1610	481,09	2206,8	10,0837	1,40909	20460
	63	0,51	790	2460	442,13	2436,1	10,7815	1,25757	9108
7	55	0,62	770	2740	171,09	2579,4	6,21074	2,68181	11946
	84	0,83	820	990	433,66	6047,2	10,3628	2,60606	12540
	51	0,54	740	1610	447,21	4127,0	5,05931	4,77272	15708
	87	0,31	490	870	509,89	2780,0	7,25750	4,34848	15642
	53	0,5	440	2510	333,71	5273,4	7,53663	4,77272	5214
	36	0,65	350	860	282,89	3840,4	9,63014	1,90909	19866
	32	0,88	550	2350	506,50	8282,7	8,33914	3,25757	9504
	71	0,46	430	1830	171,09	5187,4	7,11793	1,90909	7194
8	67	0,89	750	3090	501,42	8082,1	6,35031	3,30303	5940
	77	0,43	330	1290	225,30	5388,0	5,23377	1,46969	5280
	40	0,64	290	2840	465,85	5072,8	5,02442	3,71212	13926
	54	0,86	850	2720	354,04	4872,2	4,46615	1,68181	18612
	49	0,68	820	3130	182,95	3152,6	10,6769	2,62121	17754
	49	0,71	250	2190	409,94	2436,1	7,85066	2,54545	19536
	48	0,54	800	1670	242,24	7222,3	10,5024	1,78787	18480
	29	0,54	250	1870	154,15	7136,3	8,82763	1,68181	16962
	82	0,62	540	790	169,4	4213,0	6,55966	4,66666	19932
	53	0,76	440	1750	169,4	7938,8	8,16468	3,06060	13200
	68	0,88	820	1000	262,57	4814,8	2,89602	3,75757	6534
	36	0,69	530	2480	188,03	3295,9	5,37334	1,74242	14520
	82	0,86	550	2700	304,92	8827,2	2,96580	1,57575	12408
	82	0,86	820	2980	389,62	6161,9	7,99023	2,37878	10296
	27	0,73	410	1200	514,97	6563,1	7,67620	3,09090	8712

№ п/п	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	80	0,34	690	1380	409,94	8368,7	9,59525	4,65151	14916
9	80	0,7	520	2690	154,15	5273,4	10,3977	3,87878	11352
	75	0,39	600	1460	525,14	4213,0	4,53593	3,42424	15444
	83	0,38	490	3130	469,23	7021,7	9,24633	4,27272	6204
	38	0,64	870	890	499,73	6419,8	6,35031	4,39393	7524
	72	0,77	440	1900	398,09	4127,0	10,7117	3,33333	9240
	68	0,31	690	1370	437,05	3525,1	10,8164	3,45454	9570
	40	0,68	460	1940	269,34	3754,4	3,52407	2,74242	8052
	71	0,67	730	1400	367,59	6849,7	8,40893	3,68181	17160
	83	0,75	330	2550	479,40	7365,6	5,16399	4,74242	20394
	89	0,51	440	3110	260,87	7537,5	5,1291	3,53030	8976

7. ІНДИВІДУВАЛЬНЕ ЗАВДАННЯ № 5.

Прогнозування і методи очищення даних

Критерії оцінювання: це завдання оцінюється у 5 балів за національною шкалою. За кожну помилку знімається 0,1 бали. Потім оцінка перераховується за 100-бальною системою згідно існуючого положення.

Мета роботи: Набути навичок у визначенні точної ваги, похибок зважування та можливих втрат через це із використанням табличного процесору Excel.

- Завдання:**
1. Вибрати варіант згідно № в списку групи.
 2. За методичними матеріалами вивчити методи підготовки вихідних даних і приведення даних до стаціонарного вигляду.
 3. Вихідні данні(варіанти) методичні матеріали, приклад виконання завдання знаходиться у файлі лаборатрна_робота_5.xlsx.
 - a. Виконати підготовку даних для прогнозування;
 - b. Привести дані до стаціонарного вигляду;
 4. використати логарифмування і кінцеві різниці.
 5. Виконати прогнозування за методами:
 - А) Методом ковзної середньої для 5 точок.
 - Б) Методом експоненційного згладжування.
 - В) Методом лінійної авторегресійної функції.
 - Г) Однією із функцій серії FORECAST.
 6. Оформити звіт по Завданню в Word документі, розрахунки додати в документі Excel.