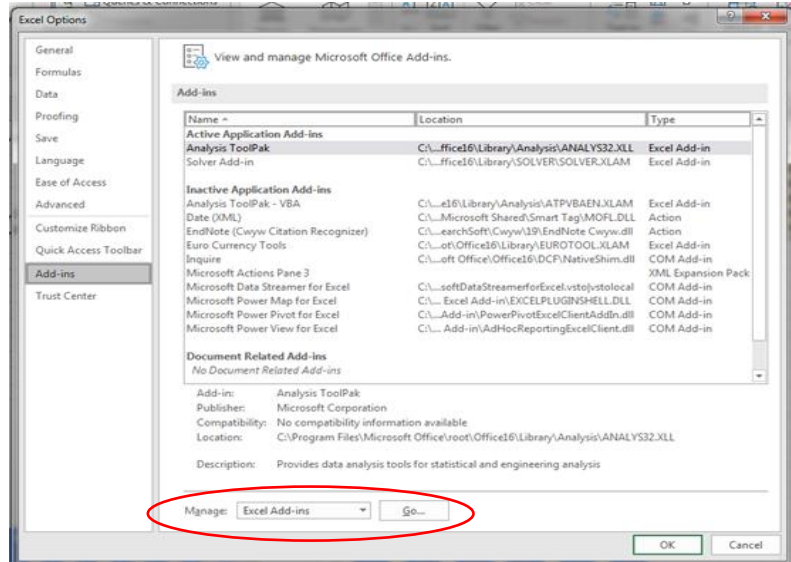


Laboratorium 2

Analysis ToolPack

Większość funkcji statystycznych w programie Excel wymaga uruchomienia dodatku Analysis ToolPack. Dodatek ten można uruchomić poprzez wybranie kolejno Plik → Opcje → Dodatki → (na dole) Dodatki programu Excel → odhaczyć Analysis ToolPack



Statystyki opisowe

Do podstawowych statystyk opisowych zaliczamy zakres danych (minimum i maksimum), średnią i odchylenie standardowe. Można je wyznaczyć za pomocą standardowych funkcji w Excelu (patrz laboratorium 1) lub za pomocą dodatku Analysis ToolPack w zakładce Dane → Analiza → Statystyki opisowe

The screenshot shows the Excel interface with the 'Data' tab selected and the 'Data Analysis' button highlighted. The 'Data Analysis' dialog box is open, showing the 'Descriptive Statistics' option selected.

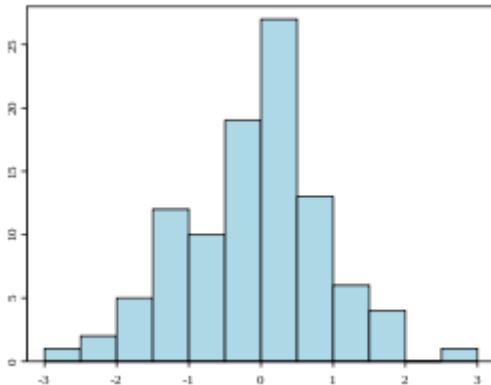
Mean	2,100661
Standard Error	0,000496
Median	2,1
Mode	2,099
Standard Deviation	0,003714
Sample Variance	1,38E-05
Kurtosis	-0,31681
Skewness	0,253744
Range	0,016
Minimum	2,094
Maximum	2,11
Sum	117,637
Count	56
Confidence Level(95,0%)	0,000995

Histogram

Histogramem nazywamy wykres kolumnowy empirycznego rozkładu pewnej cechy na podstawie próbki. Wykres przedstawia liczebność dla poszczególnych przedziałów wartości dla danych z próbki.

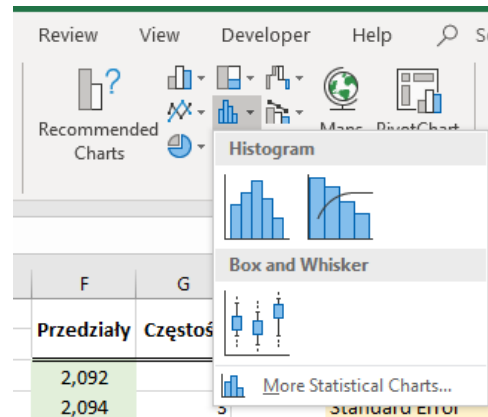
W programie Excel histogram można wstawić za pomocą dodatku Analysis ToolPack w zakładce Dane → Analiza → Histogram

W usłudze Office 365 dostępny jest też wykres Histogram automatycznie generujący histogramy z wybranych danych.



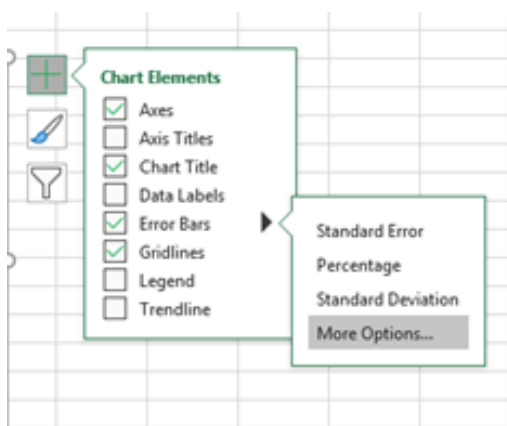
Rys 1 Przykładowy histogram

<https://pl.wikipedia.org/wiki/Histogram>



Słupki błędów na wykresach

Na wykresach umieszczamy słupki błędów, w zależności od sytuacji – błąd standardowy, odchylenie standardowe lub wartości wyznaczone eksperymentalnie. Aby dodać w programie excel słupki błędów należy kliknąć ikonę „+” przy wykresie → słupki błędów. W przypadku wyznaczonych osobno wartości wybieramy ... → więcej opcji → niestandardowe



Formuły tablicowe

Formuły tablicowe w Excelu to m.in. funkcje w wyniku których otrzymujemy wynik w wielu komórkach naraz. Aby skorzystać z formuł tablicowych należy wykonać trzy kroki:

1. Zaznaczyć wszystkie komórki wyjściowe
2. W pierwszej komórce wpisać pożądaną formułę
3. Nacisnąć Ctrl+Shift+Enter

Regresja liniowa

Regresją liniową nazywamy model statystyczny, w którym zmienna objaśniana (Y) jest opisana jako liniowa kombinacja zmiennych niezależnych (X). W najprostszym, jednowymiarowym przypadku zmienna Y jest zadana za pomocą równania linii prostej:

$$Y = aX + b$$

W programie Excel wyznaczenie równania regresji można dokonać za pomocą funkcji tablicowej REGLINP. Aby wyznaczyć parametry a i b w powyższym równaniu należy najpierw zaznaczyć dwie komórki wyjściowe, następnie wpisać formułę

$$= \text{REGLINP}(\text{zakres Y}, \text{zakres X})$$

i nacisnąć Ctrl+Shift+Enter.

Aby wyznaczyć również dodatkowe statystyki, np. błąd w wyznaczaniu parametrów a i b, należy zaznaczyć odpowiednio więcej komórek (tu: 2 na 2) i skorzystać z opcji wyświetlania dodatkowych statystyk:

$$= \text{REGLINP}(\text{zakres Y}, \text{zakres X}; \text{PRAWDA}; \text{PRAWDA})$$

PRAWDA – czy w regresji wyznaczamy parametr b

(FAŁSZ: przyjmujemy $b=0$)

PRAWDA – czy wyświetlać dodatkowe statystyki

	E	F	G	H
		X	Y	
		3,68	6,03	
		4,38	3,77	
		5,45	6,52	
		7,45	4,36	

Parametry	a	b
Y = aX + b		6,245013
Odch. Stand	0,54721	2,973082

Regresja wykładnicza

Regresją wykładniczą w przypadku jednowymiarowym nazywamy model, gdzie zmienna objaśniana Y jest modelowana za pomocą zależności:

$$Y = e^{aX} + b$$

W programie Excel do wyznaczenia równania regresji wykładniczej można wykorzystać analogiczną funkcję tablicową REGEXPP:

$$= \text{REGEXPP}(\text{zakres Y}, \text{zakres X})$$

Laboratorium 2

Jest to równoważne z wykonaniem regresji liniowej dla modelu zlinearyzowanego w postaci:

$$\ln(Y) = aX + b$$

Korelacja i R^2

W celu wyznaczenia korelacji pomiędzy dwoma zmiennymi X i Y można wykorzystać funkcję PEARSON:

$$= \text{PEARSON}(X, Y)$$

Aby wyznaczyć współczynnik R^2 dopasowania zmiennej X do zmiennej Y można wykorzystać funkcję R.KWADRAT:

$$= \text{R.KWADRAT}(Y, X)$$

Linie trendu

Do zobrazowania rozkładu jednej zmiennej względem drugiej wykorzystujemy wykres punktowy, gdzie pierwsza kolumna danych jest traktowana jako zmienna X.

Na wykresie punktowym można przedstawić przebieg krzywej regresji wykorzystując linię trendu. W tym celu należy zaznaczyć wykres punktowy, a następnie kliknąć ikonę „+” przy wykresie → linia trendu → Więcej opcji.

W pojawiającym się panelu możemy wybrać typ regresji (liniowej, wykładniczej czy innej) oraz wybrać czy na wykresie ma się automatycznie pojawić równanie krzywej oraz współczynnik R^2 .

=REGLINP(H2:H5;F2:F5;TRUE;TRUE)				
D	E	F	G	H
		X	Y	ln(Y)
		3,68	6,03	1,797441
		4,38	3,77	1,327465
		5,45	6,52	1,875405
		7,45	4,36	1,471536
Parametry		a	b	
Y = aX + b		(H2:H5;	1,811775	
Odch. Stand		0,109146	0,593009	

