Korelacija

Statistički metodi 2019/2020

dr Aleksandar Tomašević

04.12.2019.

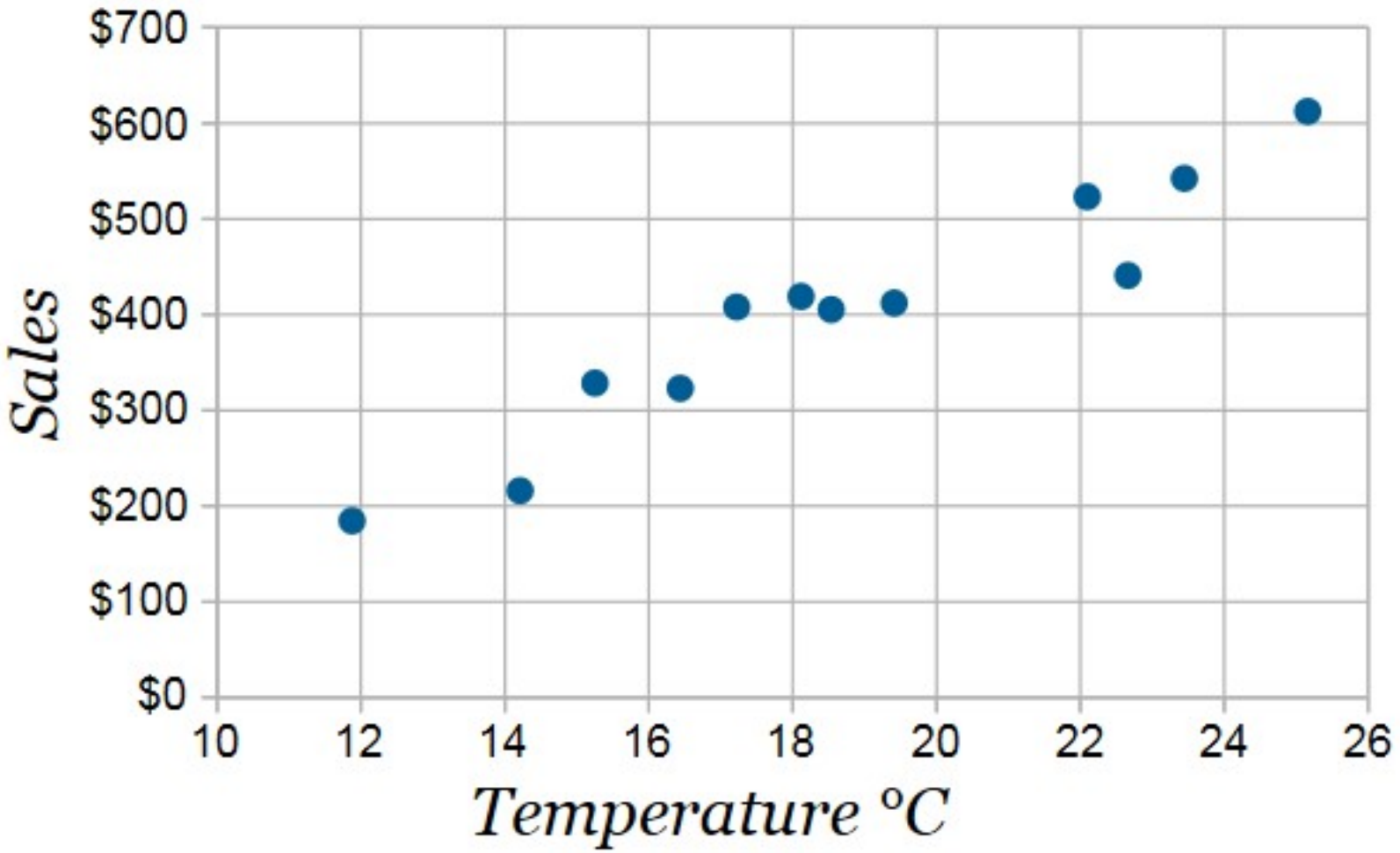
Filozofski fakultet u Novom Sadu

Plan časa

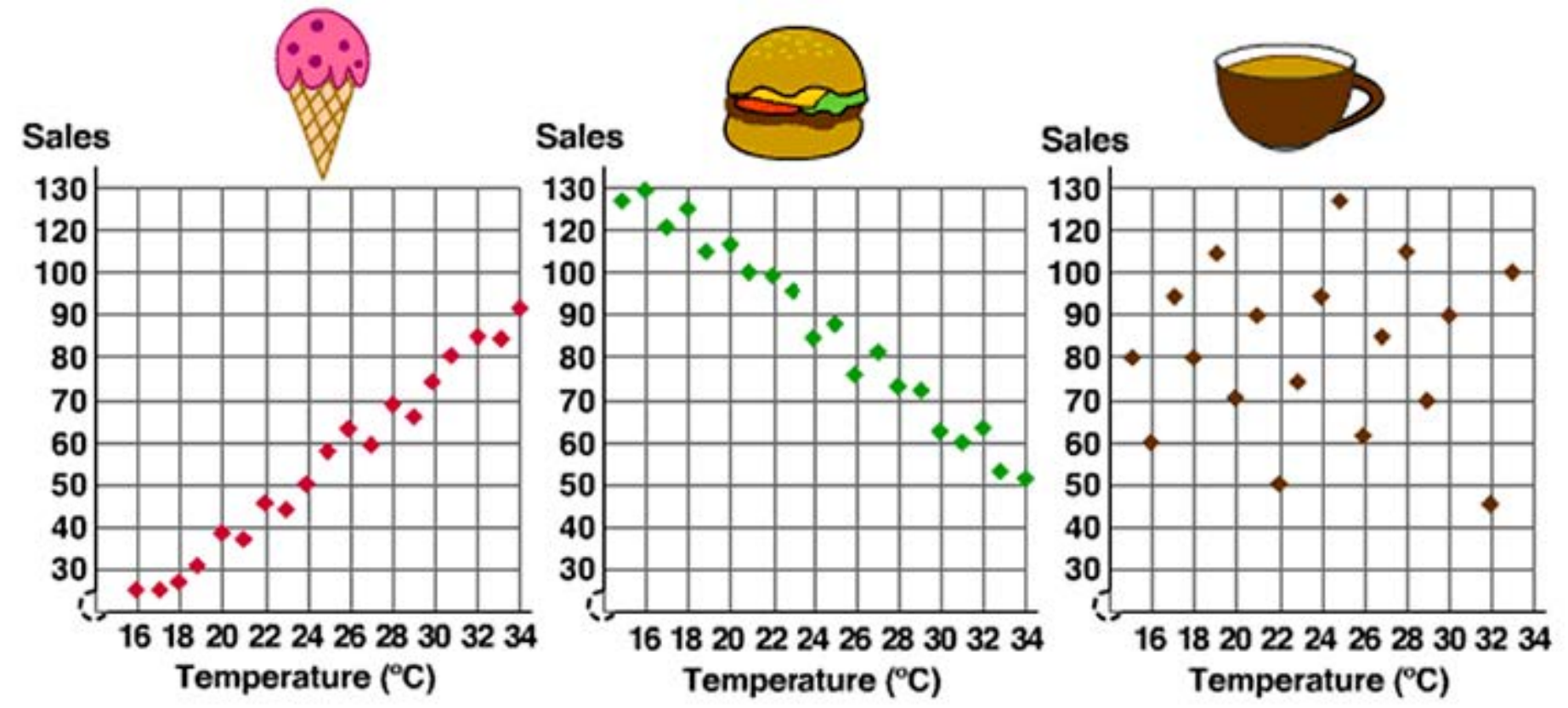
1. Dijagrami raspršenosti
2. Kovarijansa
3. Korelacija
4. Testiranje značajnosti koeficijenta korelacje

Dijagrami raspršenosti

Primer 1 - Prodaja sladoleda i temperatura



Primer 2


Sales
130
120
100
90
80
70
60
50
40
30
16 18 20 22 24 26 28 30 32 34
Temperature (°C)
Sales
130
120
100
90
80
70
60
50
40
30
16 18 20 22 24 26 28 30 32 34
Temperature (°C)
Sales
130
120
100
90
80
70
60
50
40
30
16 18 20 22 24 26 28 30 32 34
Temperature (°C)

Kovarijansa

Kovarijansa

Kovarijansa

$$cov_{XY} = \frac{\sum_{i=1}^{n}(X - \bar{X})(Y - \bar{Y})}{n - 1}$$

Korelacija

Standardni oblik formule korelacije

Korelacija

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$$

Pomoćne formule

SSX

$$SSX = \sum X^2 - \frac{(\sum X)^2}{n-1}$$

SSY

$$SSX = \sum Y^2 - \frac{(\sum Y)^2}{n-1}$$

SSP

$$SSX = \sum XY - \frac{(\sum X)(\sum Y)}{n-1}$$

Skraćeni oblik

$$r = \frac{SXY}{\sqrt{SSX * SSY}}$$

Testiranje značajnosti koeficijenta korelacje

Hipoteze

Nulta hipoteza

$$H_0 : \rho = 0$$

Alternativna hipoteza

$$H_1 : \rho \neq 0$$

T-test za korelaciju

t-statistika

$$t = \frac{r}{s_r}$$

Standardna greška korelacije

$$s_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

Primer

X	Y
0	50
1	550
2	300
2	220
5	400
6	350
6	500
8	900
9	1200
10	320
10	950