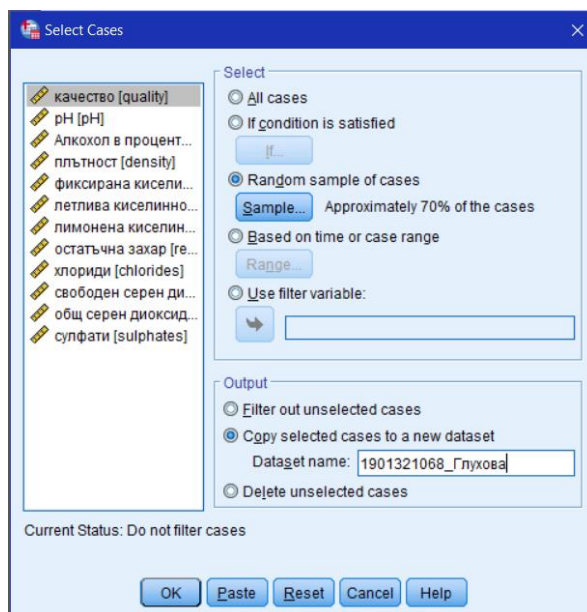


## РЕШЕНИЕ НА КУРСОВАТА ЗАДАЧА:

Направете 70% случайна извадка от файла *RED\_Wine\_4.sav*. Запишете данните в нов файл с име «ФNo\_фамилия»:

1. Сваляне на файла *RED\_Wine\_4.sav* и отварянето му в SPSS.  
**File -> Open -> Data...**
2. Генериране на случайна извадка – **70% от генералната съвкупност** и записването ѝ в нов Dataset с име *Глухова.sav*.

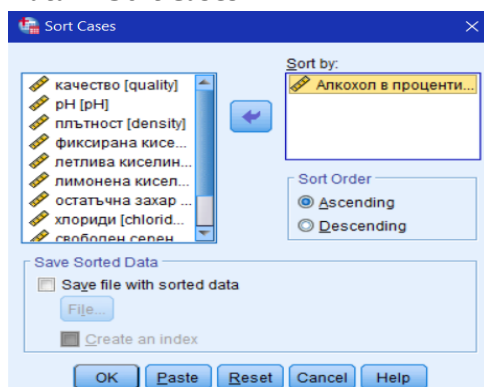
**Data -> Select Cases...**



- I. /5 т./ Направете описателна статистика променливата Алкохол на виното. Запишете мерките за централна тенденция и мерките за разсейване. Тълкувайте значението на коефициентите от гледна точка на нормалност на разпределението.

- 1) Тъй като са дадени в **негруппиран ред** ще се наложи да ги подредим по променливата алкохол във възходящ ред или както е наименованието му в статистиката – **вариационен ред**.

**Data -> Sort Cases...**



2) *Записвам мерките на алкохола:*

**Мерки на централна тенденция:**

Mean – извадково средно(средния процент на алкохол във виното) ->  $\bar{X} = 10,4696\%$

Median – медиана ->  $M_e = 10,4600$

Mode – мода ->  $M_o = 9,68$  (не е **многомодна**)

**Мерки на разсейване:**

Std. deviation – стандартно отклонение ->  $s = 1,03527$

Variance – извадкова дисперсия ->  $s^2 = 1,072$

Range – размах ->  $R = 6,52$

Minimum – най-малката стойност ->  $X_{\min} = 7,07$

Maximum – най-голямата стойност ->  $X_{\max} = 13,59$

Percentiles – квантили ->  $Q_1 = 9,7650$ ;  $M_e = Q_2 = 10,4600$ ;  $Q_3 = 11,2200$

**Коефициент на асиметрия:**  $S_k = -0.40$

Има малко повече ниски стойности на алкохола и леко струпване на данни в лявата част, но може да се каже че **разпределението е нормално**.

**Коефициент на ексцес:**  $K_u = -0,029$

Тази отрицателна стойност не е много далеч от 0-лата, така че може да се счете че **разпределението е симетрично**.

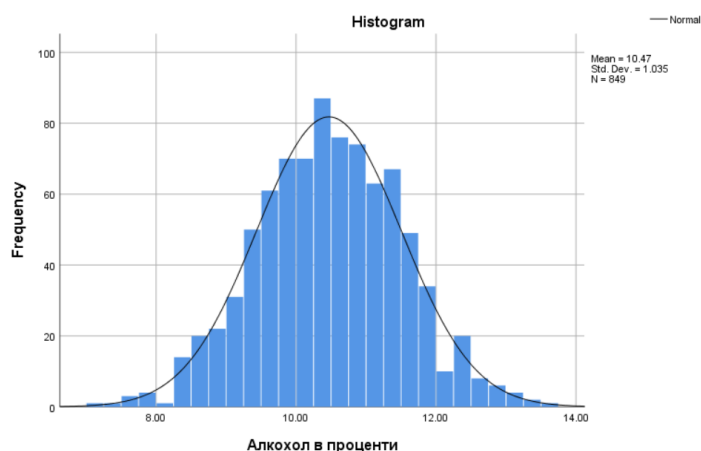
**Нормалност:**

$$\frac{|S_k|}{SE(S_k)} < p \Rightarrow 0,48 < 1,96 \quad \text{и} \quad \frac{|K_u|}{SE(K_u)} < p \Rightarrow 0,17 < 1,96$$

При ниво на значимост  $\alpha = 0.05$ ,  $p = 1.96$  **разпределението се счита за нормално**.

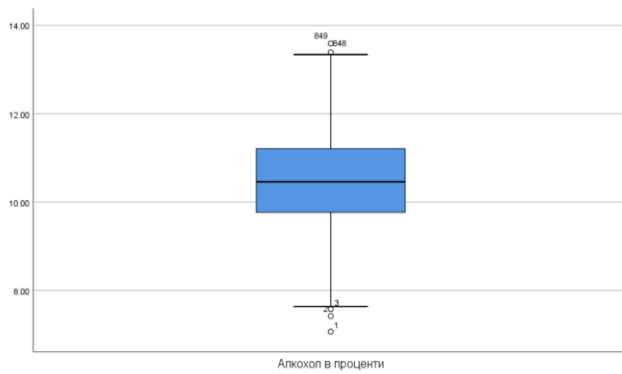
II. /5 т./ Начертайте хистограма, бокс-плот и Q-Q plot на изследваната променлива. Тълкувайте графиките от гледна точка на нормалност на разпределението.

**Хистограма:**



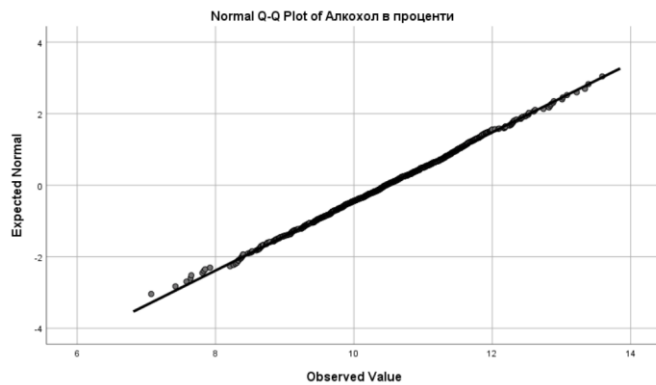
Кривата почти изпълва графиката на хистограмата, което е признак за **нормалност**.

**Бокс-плот:**



Забелязва се че **извадката е симетрична**.

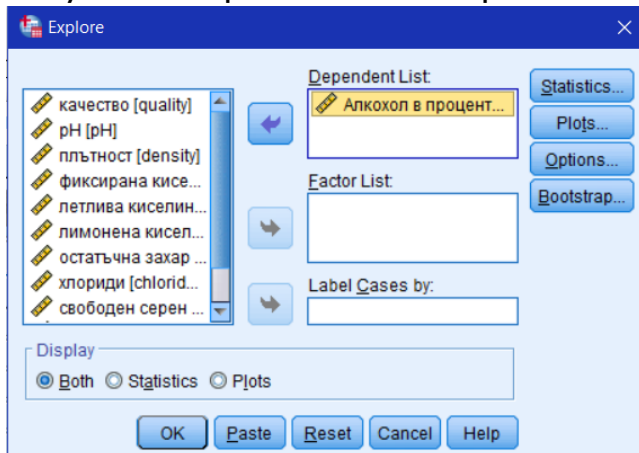
**Q-Q plot:**



Данните са близко до кривата това значи че **разпределението е нормално**.

- III. /5 т./ Направете тестове за нормалност и установете вида на разпределението на променливата Алкохол на виното.

Analyze -> Descriptive Statistics -> Explore... Plots... -> Normality plots with tests



### Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Алкохол в проценти	849	100.0%	0	0.0%	849	100.0%

От тази таблица можем да получим информация за това **колко от данните ще бъдат анализирани и колко няма, поради липсващи стойности.**

**Валидни – 100%**

**Липсващи – 0%**

### Descriptives

		Statistic	Std. Error
Алкохол в проценти	Mean	10.4696	.03553
	95% Confidence Interval for Mean	Lower Bound	10.3998
		Upper Bound	10.5393
	5% Trimmed Mean	10.4707	
	Median	10.4600	
	Variance	1.072	
	Std. Deviation	1.03527	
	Minimum	7.07	
	Maximum	13.59	
	Range	6.52	
	Interquartile Range	1.46	
	Skewness	-.040	.084
	Kurtosis	-.029	.168

**Описателната статистика съдържа мерките и коефициентите на искана статистика.**

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Алкохол в проценти	.018	849	.200 <sup>*</sup>	.999	849	.947

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

**Разпределението на променливата е близко до нормалното, защото  $p > \alpha$  ( $\text{sig} > 0.05$ ) => нямаме причина да отхвърлим нулевата хипотеза -  $H_0$ . Използвам теста за нормалност на Шапиро-Уилк, защото нямам над 2000 наблюдения.**

IV. /5 т./ Проверете хипотезата, че средната стойност на алкохолното съдържание на виното е 11 срещу алтернативната двустранна хипотеза.

1) Проверявам дали средната стойност на алкохолното съдържание на виното е 11 по следния начин:

**Analyze -> Compare Means -> One-Sample T Test...** -> Избирам съответната променлива(алкохол) и стойността 11.

Генерира ми се следното съдържание във view-a:

## T-Test

[DataSet1] C:\Users\User\Desktop\Статистика\_Даниела\_1901321068\1901321068\_Глухова.sav

### One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Алкохол в проценти	849	10.4696	1.03527	.03553

### One-Sample Test

Test Value = 11

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Алкохол в проценти	-14.928	848	.000	-.53041	-.6002	-.4607

От първата таблица разбирам че извадковото средно е 10,4696. А от втората мога да тълкувам какъв е извода от теста:

$Sig. = 0.000; \alpha = 0.05$

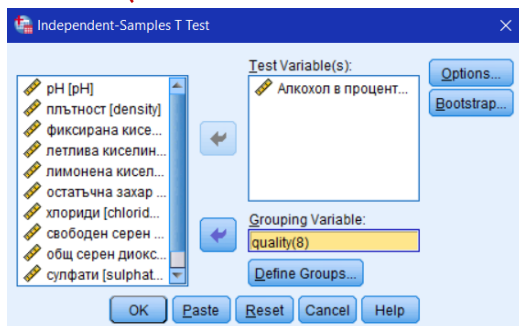
$Sig. < \alpha;$

тоест отхвърлям нулевата хипотеза  $H_0 \rightarrow \mu = 11$  и приемам алтернативната, която е  $\Rightarrow H_1 \rightarrow \mu \neq 11$ .

- V. /5 т./ Установете има ли статистическа значимо различие в алкохолното съдържание на виното при оценка за качеството под и над 8.

Analyze -> Compare Means -> Independent-Samples T Test... (за 2 независими променливи)

Избирам да тествам променливата алкохол, като я групирам по оценка на **качеството под и над 8(стойността на точката на деление)**:



Описателна статистика:

### T-Test

#### Group Statistics

	качество	N	Mean	Std. Deviation	Std. Error Mean
Алкохол в проценти	>= 8.00	46	12.0759	.61592	.09081
	< 8.00	803	10.3776	.97750	.03450

Първата група се определя от всички резултати, които са **по-големи или равни на точката на деление ( $\geq 8$ )**. Втората група се определя от всички резултати, които са **по-ниски от точката на деление ( $< 8$ )**.

## Статистика за изводи:

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Алкохол в проценти	Equal variances assumed	13.490	.000	11.648	847	.000	1.69830	.14580	1.41212	1.98447
	Equal variances not assumed			17.482	58.854	.000	1.69830	.09714	1.50390	1.89269

### 1) Тест на Levene - проверявам за равенство между дисперсии:

$H_0 \rightarrow \text{Sig}_1^2 = \text{sig}_2^2$  или  $H_1 \rightarrow \text{Sig}_1^2 \neq \text{sig}_2^2$

$\text{Sig}(p) = 0.000 < 0.05 \Rightarrow$  отхвърлям  $H_0$  (нулевата хипотеза) и приемам алтернативната  $H_1 \rightarrow \text{Sig}_1^2 \neq \text{sig}_2^2 \Rightarrow$  дисперсията на едната извадка е различна от дисперсията на другата.

Следователно попадам във втория ред (за неравенство между дисперсии).

### 2) Проверка за равни средни:

$H_0 \rightarrow \mu_1 = \mu_2$  или  $H_1 \rightarrow \mu_1 \neq \mu_2$

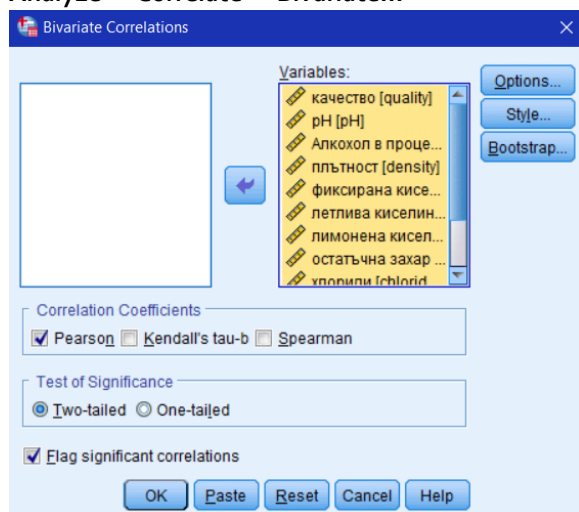
$\text{Sig}(2\text{-tailed}) = 0.000 < 0.05 \Rightarrow$  приемам алтернативната хипотеза  $H_1 \rightarrow \mu_1 \neq \mu_2$ .

$\Rightarrow$  има различие в алкохолното съдържание когато стойностите на качеството са  $\geq$  или  $<$  от 8. Алкохола в проценти за първата група (с качество  $\geq 8$ ) е значително по-голям ( $\text{Mean}_1 > \text{Mean}_2$ ) от алкохола за втора група (с качество  $< 8$ ). Качеството влияе на алкохолния процент. Използвах *effect size calculator* за да изчисля колко е ефекта на разликата ( $\text{Chen's } d = 2.07 > 0.80 \Rightarrow$  огромна разлика).

- VI. /5 т./ Изследвайте със статистически анализи наличието на възможни зависимости на алкохолното съдържание на виното от останалите променливи. Запишете всички корелационни коефициенти и тълкувайте техните знаци, стойности и тяхната значимост. Определете факторите, от които алкохолното съдържание се влияе най-силно.

Изчисляване на корелационния коефициент между данните (корелационна матрица):

### Analyze -> Correlate -> Bivariate...



Тъй като данните ми са от тип scale, ще бъдат пресмятани по метода на Pearson. Избирам също двустранен тест за значимост и да има означение за значимостта (\*  $\alpha = 0,05$ ; \*\*  $\alpha = 0,01$ ). Извадков корелационен коефициент се означава с „r“.

#### 1) Анализ за значимост :

- знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на X води до намаляване на Y)
- знака на  $r > 0$  -> пропорционална зависимост (ръста на X води до ръст на Y)
- $|r| < 0.3$  – няма корелационна зависимост
- $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- $0.7 < |r|$  ..... силна корелационна зависимост
- $\text{sig} < 0,05$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,05$
- $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$
- $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

#### Корелационна матрица (12x12):

Correlations												
	качество	pH	Алкохол в проценти	плътност	фиксирана киселинност	летлива киселинност	лимонена киселина	остатъчна захар	хлориди	свободен серен диоксид	общ серен диоксид	сулфати
качество	Pearson Correlation	1	.064	.729**	-.343**	.028	-.367**	.174**	-.007	-.227**	-.057	-.202**
	Sig. (2-tailed)		.061	.000	.000	.424	.000	.000	.839	.000	.098	.000
	N	849	849	849	849	849	849	849	849	849	849	849
pH	Pearson Correlation	.064	1	.221**	-.324**	-.684**	.244**	-.560**	-.091**	-.278**	.086*	-.054
	Sig. (2-tailed)			.000	.000	.000	.000	.008	.000	.012	.114	.000
	N	849	849	849	849	849	849	849	849	849	849	849
Алкохол в проценти	Pearson Correlation	.729**	.221**	1	-.509**	-.089**	-.178**	.046	-.036	-.264**	-.054	-.204**
	Sig. (2-tailed)				.000	.010	.000	.176	.300	.000	.116	.000
	N	849	849	849	849	849	849	849	849	849	849	849
плътност	Pearson Correlation	-.343**	-.324**	-.509**	1	.655**	.023	.367**	.374**	.238**	.000	.093**
	Sig. (2-tailed)					.000	.500	.000	.000	.000	.990	.007
	N	849	849	849	849	849	849	849	849	849	849	849
фиксирана киселинност	Pearson Correlation	.028	-.684**	-.089**	.655**	1	-.270**	.680**	.121**	.108**	-.160**	-.116**
	Sig. (2-tailed)		.424	.000	.010	.000	.000	.000	.002	.000	.001	.000
	N	849	849	849	849	849	849	849	849	849	849	849
летлива киселинност	Pearson Correlation	-.367**	.244**	-.178**	.023	-.270**	1	-.583**	.012	.068*	.019	.099**
	Sig. (2-tailed)		.000	.000	.500	.000	.000	.000	.721	.048	.584	.004
	N	849	849	849	849	849	849	849	849	849	849	849
лимонена киселина	Pearson Correlation	.174**	-.560**	.046	.367**	.680**	-.583**	1	.133**	.211**	-.069*	.015
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.045	.656
	N	849	849	849	849	849	849	849	849	849	849	849
остатъчна захар	Pearson Correlation	-.007	-.091**	-.036	.374**	.121**	.133**	1	.106**	.248**	.207**	.023
	Sig. (2-tailed)		.839	.008	.300	.000	.000	.721	.000	.000	.000	.509
	N	849	849	849	849	849	849	849	849	849	849	849
хлориди	Pearson Correlation	-.227**	-.278**	-.264**	.238**	.108**	.068*	.211**	1	.027	.046	.374**
	Sig. (2-tailed)		.000	.000	.000	.002	.048	.000	.002		.436	.181
	N	849	849	849	849	849	849	849	849	849	849	849
свободен серен диоксид	Pearson Correlation	-.057	.086*	-.054	.000	-.160**	.019	-.069*	.248**	1	.663**	.039
	Sig. (2-tailed)		.098	.012	.116	.990	.000	.584	.045	.000	.436	.000
	N	849	849	849	849	849	849	849	849	849	849	849
общ серен диоксид	Pearson Correlation	-.202**	-.054	-.204**	.093**	-.116**	.099**	.015	.207**	.046	.663**	1
	Sig. (2-tailed)		.000	.114	.000	.007	.001	.004	.656	.000	.181	.000
	N	849	849	849	849	849	849	849	849	849	849	849
сулфати	Pearson Correlation	.200**	-.216**	.073*	.157**	.184**	-.268**	.316**	.023	.374**	.039	.037
	Sig. (2-tailed)		.000	.000	.034	.000	.000	.000	.509	.000	.262	.287
	N	849	849	849	849	849	849	849	849	849	849	849

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

#### 2) Корелационни знаци, стойности и значимост:

Корелация между pH и качество ->  $r = 0,064$ ;  $\text{sig} = 0,061$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на pH води до ръст на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между алкохол в проценти и качество ->  $r = 0,729^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на алкохол в проценти води до ръст на качество)
- ❖  $0.7 < |r|$  ..... силна корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между алкохол в проценти и pH ->  $r = 0,221^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на алкохол в проценти води до ръст на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между плътност и качество ->  $r = -0,343^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на плътност води до намаляване на качество)
- ❖  $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между плътност и pH ->  $r = -0,324^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на плътност води до намаляване на pH)
- ❖  $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между плътност и алкохол в проценти ->  $r = -0,509^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на плътност води до намаляване на алкохол в проценти)
- ❖  $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между фиксирана киселинност и качество ->  $r = 0,028$ ;  $\text{sig} = 0,424$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на фиксирана киселинност води до ръст на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между фиксирана киселинност и pH ->  $r = -0,684^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на фиксирана киселинност води до намаляване на pH)
- ❖  $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между фиксирана киселинност и алкохол в проценти ->  $r = -0,089^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,10$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на фиксирана киселинност води до намаляване на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между фиксирана киселинност и плътност ->  $r = 0,665^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на фиксирана киселинност води до ръст на плътност)



- ❖  $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- ❖  $\text{sig} < 0,01$  - > корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между летлива киселинност и качество ->  $r = -0,367^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на летлива киселинност води до намаляване на качество)
- ❖  $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- ❖  $\text{sig} < 0,01$  - > корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между летлива киселинност и pH ->  $r = 0,244^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на летлива киселинност води до ръст на pH)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  - > корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между летлива киселинност и алкохол в проценти ->  $r = -0,178^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на летлива киселинност води до намаляване на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  - > корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между летлива киселинност и плътност ->  $r = 0,023$ ;  $\text{sig} = 0,500$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на летлива киселинност води до ръст на плътност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  - > корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между летлива киселинност и фиксирана киселинност ->  $r = -0,270^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на летлива киселинност води до намаляване на фиксирана киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  - > корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между лимонена киселина и качество ->  $r = 0,174^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на лимонена киселина води до ръст на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  - > корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между лимонена киселина и pH ->  $r = -0,560^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на лимонена киселина води до намаляване на pH)
- ❖  $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- ❖  $\text{sig} < 0,01$  - > корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между лимонена киселина и алкохол в проценти ->  $r = 0,046$ ;  $\text{sig} = 0,176$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на лимонена киселина води до ръст на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  - > корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между лимонена киселина и плътност ->  $r = 0,367^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на лимонена киселина води до ръст на плътност)
- ❖  $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между лимонена киселина и фиксирана киселинност ->  $r = 0,680^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на лимонена киселина води до ръст на фиксирана киселинност)
- ❖  $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между лимонена киселина и летлива киселинност ->  $r = -0,583^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на лимонена киселина води до намаляване на летлива киселинност)
- ❖  $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между остатъчна захар и качество ->  $r = -0,007$ ;  $\text{sig} = 0,839$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на остатъчна захар води до намаляване на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между остатъчна захар и pH ->  $r = -0,091^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,008$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на остатъчна захар води до намаляване на pH)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между остатъчна захар и алкохол в проценти ->  $r = -0,036$ ;  $\text{sig} = 0,300$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на остатъчна захар води до намаляване на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между остатъчна захар и плътност ->  $r = 0,374^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на остатъчна захар води до ръст на плътност)
- ❖  $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между остатъчна захар и фиксирана киселинност ->  $r = 0,121^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на остатъчна захар води до ръст на фиксирана киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между остатъчна захар и летлива киселинност ->  $r = 0,012$ ;  $\text{sig} = 0,721$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на остатъчна захар води до ръст на летлива киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между остатъчна захар и лимонена киселина ->  $r = 0,133^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на остатъчна захар води до ръст на лимонена киселина)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между хлориди и качество ->  $r = -0,227^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на хлориди води до намаляване на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между хлориди и pH ->  $r = -0,278^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на хлориди води до намаляване на pH)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между хлориди и алкохол в проценти ->  $r = -0,264^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на хлориди води до намаляване на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между хлориди и плътност ->  $r = 0,238^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на хлориди води до ръст на плътност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между хлориди и фиксирана киселинност ->  $r = 0,108^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,002$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на хлориди води до ръст на фиксирана киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между хлориди и летлива киселинност ->  $r = 0,068^{*}$  ( $\alpha = 0,05$ ) ;  $\text{sig} = 0,048$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на хлориди води до ръст на летлива киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между хлориди и лимонена киселина ->  $r = 0,211^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на хлориди води до ръст на лимонена киселина)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между хлориди и остатъчна захар ->  $r = 0,106^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,002$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на хлориди води до ръст на остатъчна захар)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между свободен сериен диоксид и качество ->  $r = -0,057$ ;  $\text{sig} = 0,098$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на свободен сериен диоксид води до намаляване на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между свободен сериен диоксид и pH ->  $r = 0,086^*$  ( $\alpha = 0,05$ );  $\text{sig} = 0,012$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на свободен сериен диоксид води до ръст на pH)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,05$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,05$

Корелация между свободен сериен диоксид и алкохол в проценти ->  $r = -0,054$ ;  $\text{sig} = 0,116$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на свободен сериен диоксид води до намаляване на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между свободен сериен диоксид и плътност ->  $r = 0,000$ ;  $\text{sig} = 0,990$

- ❖  $r = 0$  -> **НЯМА КОРЕЛАЦИЯ**
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$
- ❖  $|r| < 0.3$  – няма корелационна зависимост

Корелация между свободен сериен диоксид и фиксирана киселинност ->  $r = -0,160^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на свободен сериен диоксид води до намаляване на фиксирана киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между свободен сериен диоксид и летлива киселинност ->  $r = 0,019$ ;  $\text{sig} = 0,584$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на свободен сериен диоксид води до ръст на летлива киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между свободен сериен диоксид и лимонена киселина ->  $r = -0,069^*$  ( $\alpha = 0,05$ );  $\text{sig} = 0,045$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на свободен сериен диоксид води до намаляване на лимонена киселина)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,05$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,05$

Корелация между свободен сериен диоксид и остатъчна захар ->  $r = 0,248^{**}$  ( $\alpha = 0,01$ );  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на свободен сериен диоксид води до ръст на остатъчна захар)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между свободен сериен диоксид и хлориди ->  $r = 0,027$ ;  $\text{sig} = 0,436$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на свободен сериен диоксид води до ръст на хлориди)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между общ сериен диоксид и качество ->  $r = -0,202^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на общ сериен диоксид води до намаляване на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между общ сериен диоксид и pH ->  $r = -0,054$ ;  $\text{sig} = 0,114$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на общ сериен диоксид води до намаляване на pH)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между общ сериен диоксид и алкохол в проценти ->  $r = -0,204^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на общ сериен диоксид води до намаляване на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между общ сериен диоксид и плътност ->  $r = 0,093^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,007$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на общ сериен диоксид води до ръст на плътност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между общ сериен диоксид и фиксирана киселинност ->  $r = -0,116^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,001$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на общ сериен диоксид води до намаляване на фиксирана киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между общ сериен диоксид и летлива киселинност ->  $r = 0,099^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,004$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на общ сериен диоксид води до ръст на летлива киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между общ сериен диоксид и лимонена киселина ->  $r = 0,015$ ;  $\text{sig} = 0,656$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на общ сериен диоксид води до ръст на лимонена киселина)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между общ сериен диоксид и остатъчна захар ->  $r = 0,207^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на общ сериен диоксид води до ръст на остатъчна захар)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между общ сериен диоксид и хлориди ->  $r = 0,046$ ;  $\text{sig} = 0,181$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на общ сериен диоксид води до ръст на хлориди)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между общ сериен диоксид и свободен сериен диоксид ->  $r = 0,663^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на общ сериен диоксид води до ръст на свободен сериен диоксид)
- ❖  $0.5 < |r| < 0.7$  ... умерена корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и качество ->  $r = 0,200^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на качество)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и рН ->  $r = -0,216^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на сулфати води до намаляване на рН)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и алкохол в проценти ->  $r = 0,073^{*}$  ( $\alpha = 0,05$ ) ;  $\text{sig} = 0,034$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на алкохол в проценти)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,05$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,05$

Корелация между сулфати и плътност ->  $r = 0,157^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на плътност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и фиксирана киселинност ->  $r = 0,184^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на фиксирана киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и летлива киселинност ->  $r = -0,268^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r < 0$  -> обратно пропорционална зависимост (ръста на сулфати води до намаляване на летлива киселинност)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и лимонена киселина ->  $r = 0,316^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на лимонена киселина)
- ❖  $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и остатъчна захар ->  $r = 0,023$ ;  $\text{sig} = 0,509$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на остатъчна захар)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между сулфати и хлориди ->  $r = 0,374^{**}$  ( $\alpha = 0,01$ ) ;  $\text{sig} = 0,000$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на хлориди)
- ❖  $0.3 < |r| < 0.5$  ... слаба корелационна зависимост
- ❖  $\text{sig} < 0,01$  -> корелационния коефициент е значим при ниво на доверие  $\alpha = 0,01$

Корелация между сулфати и свободен сериен диоксид ->  $r = 0,039$ ;  $\text{sig} = 0,262$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на свободен сериен диоксид)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

Корелация между сулфати и общ сериен диоксид ->  $r = 0,037$ ;  $\text{sig} = 0,287$

- ❖ знака на  $r > 0$  -> пропорционална зависимост (ръста на сулфати води до ръст на общ сериен диоксид)
- ❖  $|r| < 0.3$  – няма корелационна зависимост
- ❖  $\text{sig} \geq 0,05$  -> корелационния коефициент е незначим при ниво на доверие  $\alpha = 0,05$

**Извод: Най - силно алкохолното съдържание се влияе от: качеството и плътността.**

VII. /5 т./ Постройте адекватен многомерен линеен регресионен модел за алкохолното съдържание на виното, като за предиктори изберете всички останали променливи. Опишете регресионния модел детерминация, адекватност на модела, значимост на коефициентите и др. Интерпретирайте получените резултати.

Analyze -> Regression -> Linear...

Таблица, която описва кои променливи участват:

Model	Variables Entered	Variables Removed	Method
1	сулфати, остатъчна захар, качество, фиксирана киселинност, свободен серен диоксид, летлива киселинност, хлориди, общ серен диоксид, pH, лимонена киселина, плътност <sup>b</sup>	.	Enter

a. Dependent Variable: Алкохол в проценти

b. All requested variables entered.

Таблица на коефициентите на корелация и детерминация:

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.850 <sup>a</sup>	.722	.719	.54926

a. Predictors: (Constant), сулфати, остатъчна захар, качество, фиксирана киселинност, свободен серен диоксид, летлива киселинност, хлориди, общ серен диоксид, pH, лимонена киселина, плътност

➤ **R** – коефициент на множествената корелация (Pearson)



- $R^2$  – коефициент на детерминация ->  $R^2 = 72,2\%$  (процента на дисперсията в зависимата променлива, който може да се обясни с независимите променливи)
  - $R^2 = \text{Sum of Squares(Regression)} / \text{Sum of Squares(Total)}$
- Adjusted  $R^2$  – подобрения коефициент

**Извод 1: Независимите променливи обясняват 72,2% от променливостта на зависимата променлива.**

Анализ на дисперсията:

ANOVA <sup>a</sup>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	656.361	11	59.669	197.784	.000 <sup>b</sup>
	Residual	252.514	837	.302		
	Total	908.875	848			

a. Dependent Variable: Алкохол в проценти

b. Predictors: (Constant), сулфати, остатъчна захар, качество, фиксирана киселинност, свободен серен диоксид, летлива киселинност, хлориди, общ серен диоксид, рН, лимонена киселина, плътност

- F - тества дали общият регресионен модел е подходящ за данните ( $\text{Sig.}(p) < 0.0005$ )

**Извод 2: Регресионния модел е подходящ за данните (статистически значим).**

Коефициенти:

Coefficients <sup>a</sup>					
Model		Unstandardized Coefficients		Standardized Coefficients	Sig.
		B	Std. Error	Beta	
1	(Constant)	361.541	17.445		.000
	качество	.778	.043	.448	.000
	рН	2.853	.207	.412	.000
	плътност	-371.659	17.844	-.729	.000
	фиксирана киселинност	.360	.027	.589	.000
	летлива киселинност	.832	.149	.143	.000
	лимонена киселина	.521	.173	.099	.003
	остатъчна захар	.131	.014	.196	.000
	хлориди	-.485	.489	-.023	.322
	свободен серен диоксид	-.001	.002	-.015	.548
	общ серен диоксид	.000	.001	-.003	.897
	сулфати	.546	.133	.090	.000

a. Dependent Variable: Алкохол в проценти



- Unstandardized Coefficients (B) – резултата от регресионното уравнение ( $\hat{y} = a_1 * x - a_0$ )
- Standardized Coefficients Beta – вероятността на факторите
  - *Увеличаването на плътността води до намаляване на общата детерминация (тъй като най- влиятелния фактор – плътността е с отрицателен знак).*
- t – статистики на константите в модела (получава се когато разделим B на Std. Error)
- Sig. – стойността на p, която ми помага да определя дали коефициента е значим и има влияние върху модела (ако е незначим може да бъде премахнат с цел подобряване на детерминацията)

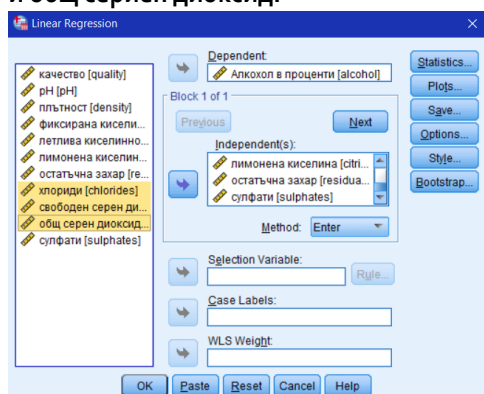
Проверка за значимост на корелационния коефициент:

- $H_0 : a = 0$  или алтернативната хипотеза  $H_1 : a \neq 0$  (която ме интересува при доказване тоест трябва да се заяви че има промяна/влияние върху модела)
  - $Sig.(p) < \alpha \rightarrow$  коефициента е значим
  - $Sig.(p) > \alpha \rightarrow$  коефициента е незначим
- 
- ❖ Корелационен коефициент  $a_0 = 0,778 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_1 = 361,541 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_2 = 2,853 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_3 = -371,659 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_4 = 0,360 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_5 = 0,832 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_6 = 0,521 \neq 0$ ;  $Sig.(0,003) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_7 = 0,131 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.
  - ❖ Корелационен коефициент  $a_8 = -0,485 \neq 0$ ;  $Sig.(0,322) > \alpha \Rightarrow$  **коефициента е значим, но няма влияние върху модела.**
  - ❖ Корелационен коефициент  $a_9 = -0,001 \neq 0$ ;  $Sig.(0,548) > \alpha \Rightarrow$  **коефициента е незначим и няма влияние върху модела.**
  - ❖ Корелационен коефициент  $a_{10} = 0,000 = 0$ ;  $Sig.(0,897) < \alpha \Rightarrow$  **коефициента е незначим и няма влияние върху модела.**
  - ❖ Корелационен коефициент  $a_{11} = 0,546 \neq 0$ ;  $Sig.(0,000) < \alpha \Rightarrow$  коефициента е значим и има влияние върху модела.

**Извод 3: Ще е добре да се премахнат факторите – хлориди, свободен сериен диоксид и общ сериен диоксид, тъй като нямат влияние върху модела и това ги прави незначими.**

VIII. /5 т./ Постройте нов многомерен линеен регресионен модел за алкохолното съдържание на виното, като отстраните незначимите предиктори, установени в модела от т.7, ако има такива. Опишете показателите на новия регресионен модел. Коментирайте го - подобриха ли се?

Отново повтарям: Analyze -> Regression -> Linear... ще отстраня - хлориди, свободен сериен диоксид и общ сериен диоксид.



**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.849 <sup>a</sup>	.722	.719	.54886

a. Predictors: (Constant), фиксирана киселинност, качество, остатъчна захар, сулфати, летлива киселинност, pH, лимонена киселина, плътност

- Има намаление на коефициента на множествената корелация и стандартното отклонение на грешката.
- Няма разлика в процента на дисперсията на зависимата променлива, който може да се обясни с независимите променливи.
- Няма разлика в подобрения коефициент.

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	655.830	8	81.979	272.133	.000 <sup>b</sup>
	Residual	253.046	840	.301		
	Total	908.875	848			

a. Dependent Variable: Алкохол в проценти

b. Predictors: (Constant), фиксирана киселинност, качество, остатъчна захар, сулфати, летлива киселинност, pH, лимонена киселина, плътност

- Има увеличение в стойностите на регресията, тоест има статистически разлики от промяната.

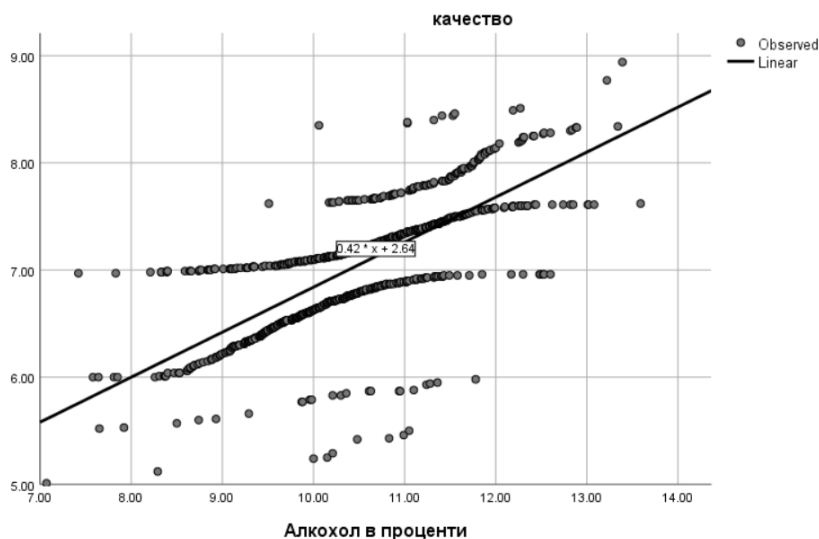
Coefficients <sup>a</sup>					
Model		Unstandardized Coefficients		Standardized Coefficients	Sig.
		B	Std. Error	Beta	
1	(Constant)	363.802	17.109		21.264
	качество	.788	.042	.453	18.751
	pH	2.909	.194	.420	14.976
	плътност	-374.258	17.441	-.734	21.459
	летлива киселинност	.802	.142	.138	5.645
	лимонена киселина	.476	.163	.091	2.917
	сулфати	.490	.123	.081	3.970
	остатъчна захар	.128	.014	.192	9.092
	фиксирана киселинност	.369	.025	.605	14.715

a. Dependent Variable: Алкохол в проценти

➤ При факторите вероятността се е увеличила, но все още имам обратна пропорционална зависимост.

IX. /5 т./ Запазете предсказаните данни от модела и запишете уравнението на регресия, което е с по-добри показатели.

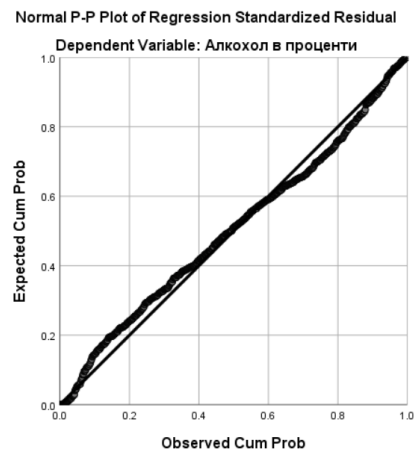
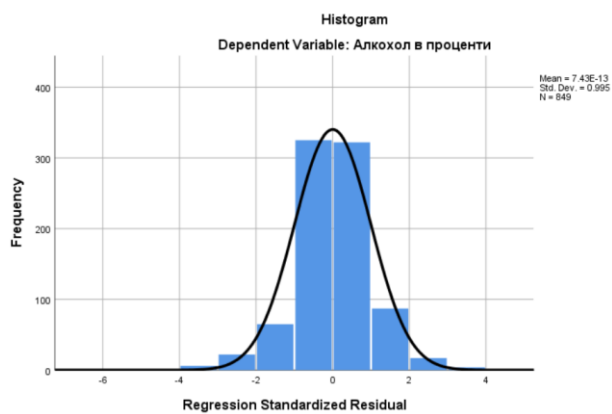
Analyze -> Regression -> Curve Estimation



Качество:  $\hat{y} = 0.42 * x + 2.64$

X. /5 т./ Изследвайте разпределението на остатъците с тестове и графики. Направете изводи.

Analyze -> Regression -> Linear... -> Save... (Predicted Values -> Unstandardized ;Residuals -> Unstandardized) и Plots... (Histogram и Normal probability plot; ZPRED – X; ZPRESID - Y) – за остатъците



- ✓ Кривата почти изпълва графиката на хистограмата, което е признак за **нормалност**.
- ✓ Данните са близко до кривата това значи че **разпределението е нормално**.

**Извод: Разпределението на остатъците е нормално.**