

VERİ SETİ:

kaza_sayısı	bosanma_sayisi	Evli_Mutlu	Evli_degil_Mutlu	Evli_Nötr	Evli_degil_Nötr	Evli_Mutsuz	Evli_degil_Mutsuz	Mutlu	Nötr	Mutsuz	Evlenme_Yası
23963	4332	56,44	46,11	31,28	34,8	12,27	19,09	53	32,45	14,54	27
2645	551	67,05	60,65	23,37	24,69	9,58	14,66	65,01	23,79	11,2	26
1441	252	62,23	54,18	23,23	24,2	14,54	21,62	60,09	23,49	16,42	24
3270	463	68,37	57,57	25,73	32,16	5,9	10,27	66,02	27,13	6,85	26
103051	11195	58,44	50,18	33	36,06	8,56	13,75	56,23	33,82	9,95	27
:	:	:	:	:	:	:	:	:	:	:	:
1263	158	66,43	53,73	25,78	32,85	7,79	13,42	62,72	27,85	9,43	28
12272	2330	56,33	46,93	34,01	36,47	9,67	16,6	53,46	34,76	11,78	27
13411	2416	72,59	65,84	21,13	23,69	6,28	10,47	70,73	21,83	7,43	26
2078	389	64,85	60,05	29,45	33,97	5,7	5,98	63,55	30,67	5,78	26
1265	162	61,3	55,45	26,85	27,81	11,85	16,74	59,49	27,15	13,36	27

Veri seti üstünde **işlemler** yapılması için veriler hakkında **bilgi sahibi** olunması gerekir.

DATA COLUMNS (TOTAL 12 COLUMNS):

COLUMN	NON-NULL COUNT	DTYPE
KAZA_SAYISI	81 NON-NULL	INT64
BOSANMA_SAYISI	81 NON-NULL	INT64
EVLI_MUTLU	81 NON-NULL	FLOAT64
EVLI_DEGIL_MUTLU	81 NON-NULL	FLOAT64
EVLI_NÖTR	81 NON-NULL	FLOAT64
EVLI_DEGIL_NÖTR	81 NON-NULL	FLOAT64
EVLI_MUTSUZ	81 NON-NULL	FLOAT64
EVLI_DEGIL_MUTSUZ	81 NON-NULL	FLOAT64
MUTLU	81 NON-NULL	FLOAT64
NÖTR	81 NON-NULL	FLOAT64
MUTSUZ	81 NON-NULL	FLOAT64
EVLENME_YASI	81 NON-NULL	INT64

DYPES: FLOAT64(9), INT64(3)

Veri üzerinde istatistiksel işlemler yapılabilmesi için bazı sütunların veri tipleri float64 e çevrilir.

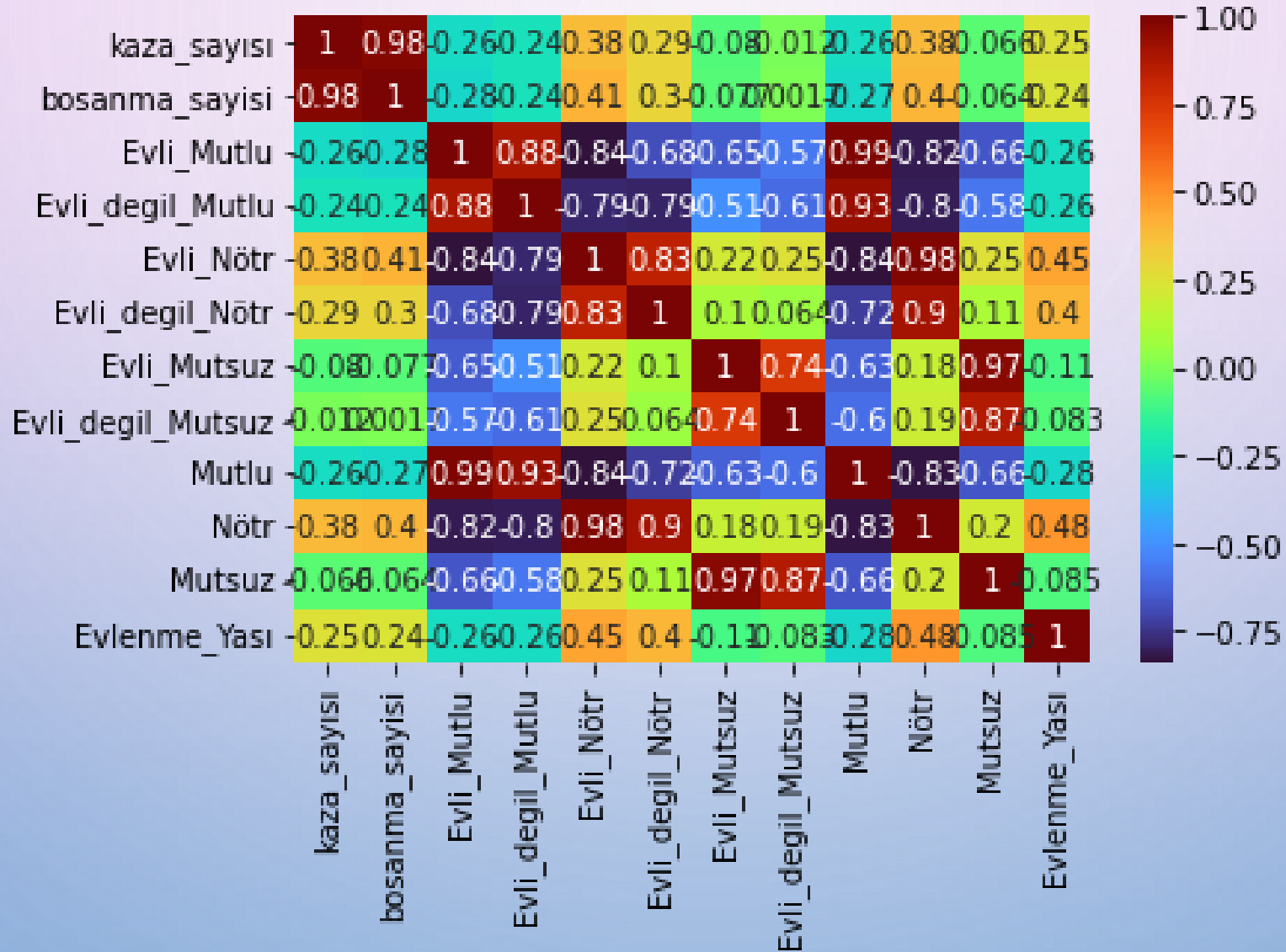
kaza_sayısı	:	0.0
bosanma_sayisi	:	0.0
Evli_Mutlu	:	0.6948
Evli_degil_Mutlu	:	0.1736
Evli_Nötr	:	0.1331
Evli_degil_Nötr	:	0.0114
Evli_Mutsuz	:	0.0
Evli_degil_Mutsuz	:	0.0182
Mutlu	:	0.535
Nötr	:	0.1289
Mutsuz	:	0.0
Evlenme_Yası	:	0.0

Verinin normal dağılıp dağılmadığını bulmak için **Shapiro Wilk** testi uygulanır ve p değeri $\alpha=0,05$ değeri ile karşılaştırılır.

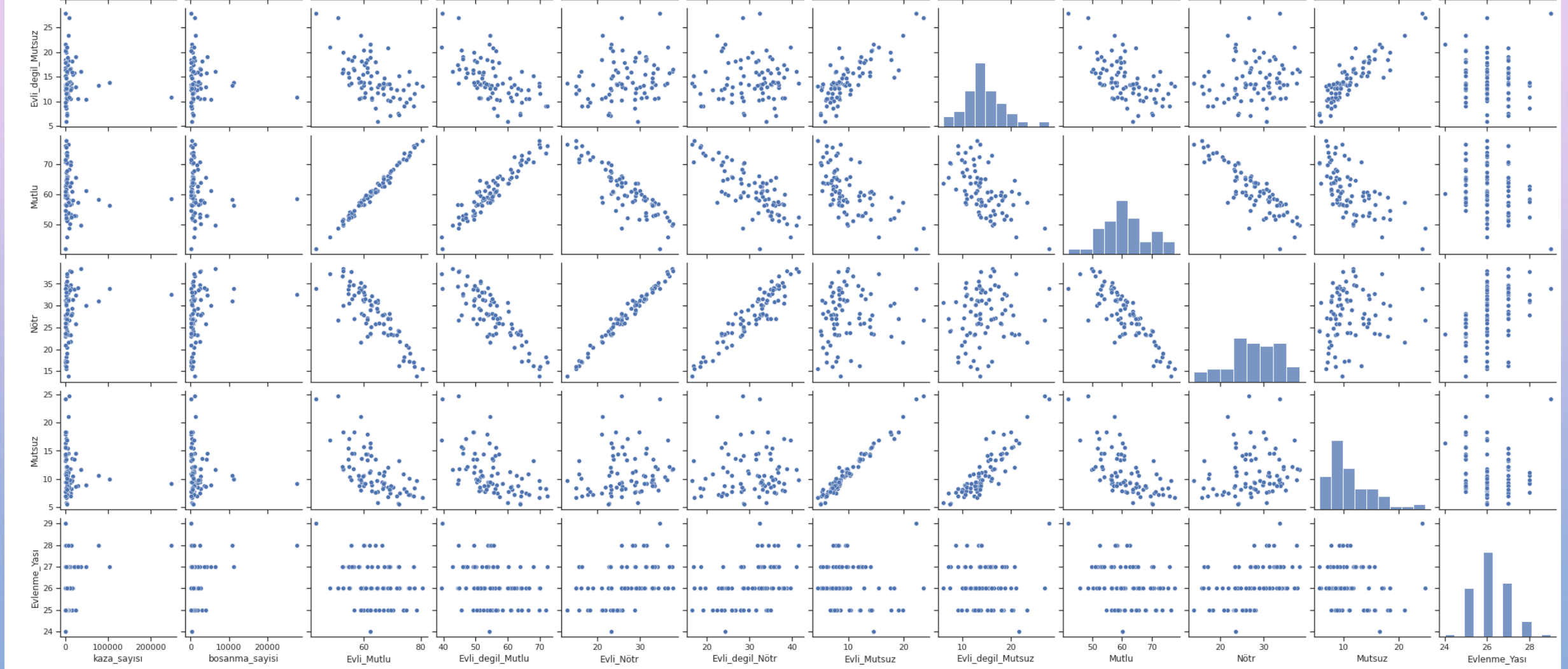
Tabloda bulunan sütunların genel olarak normal dağılmadığını görmekteyiz.

Verilerin normalliği kullanılacak korelasyon testinin seçimi için önemlidir.

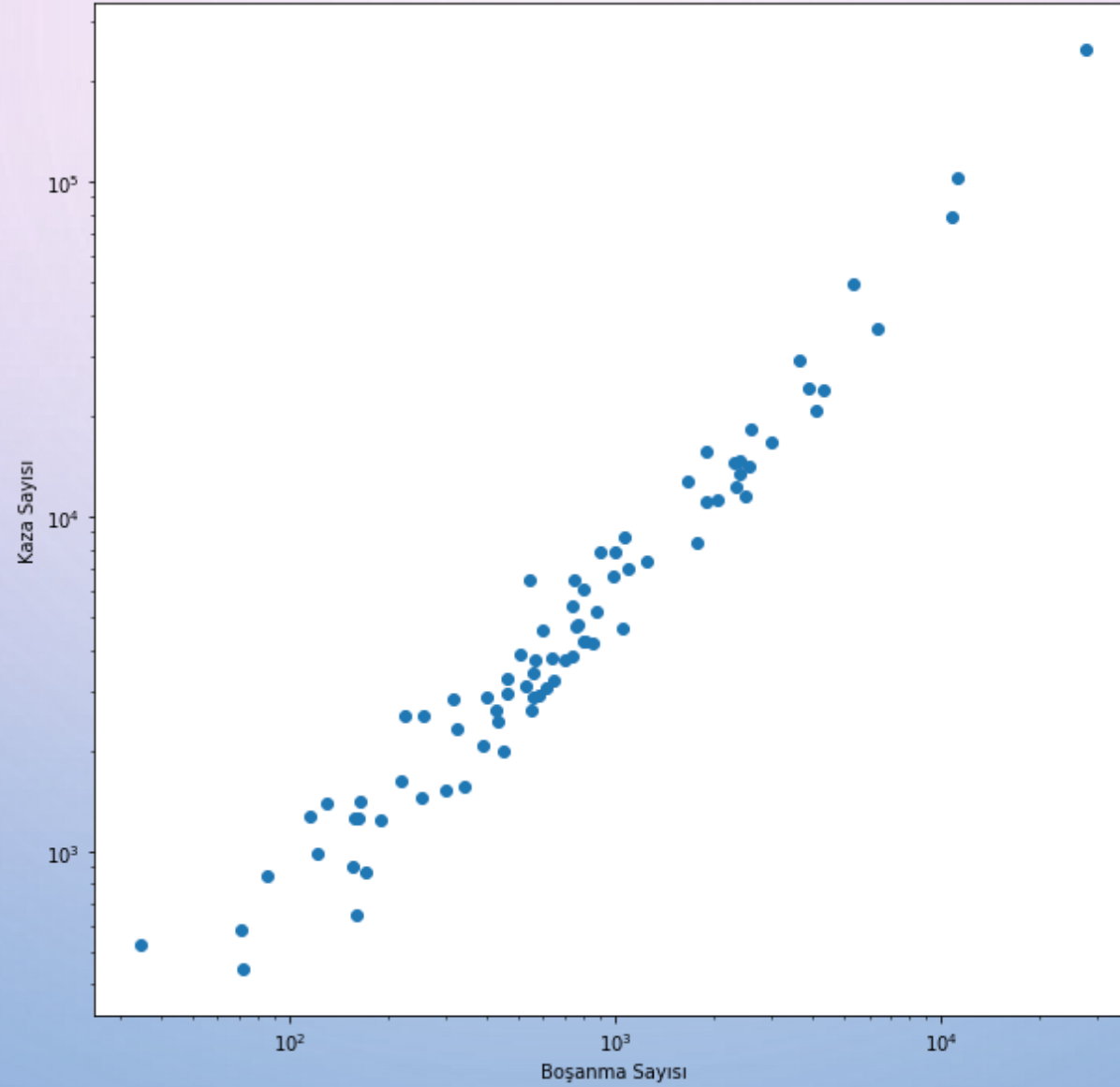
Veriler normal dağılmadığı için **Spearman** korelasyon testi uygulanır.



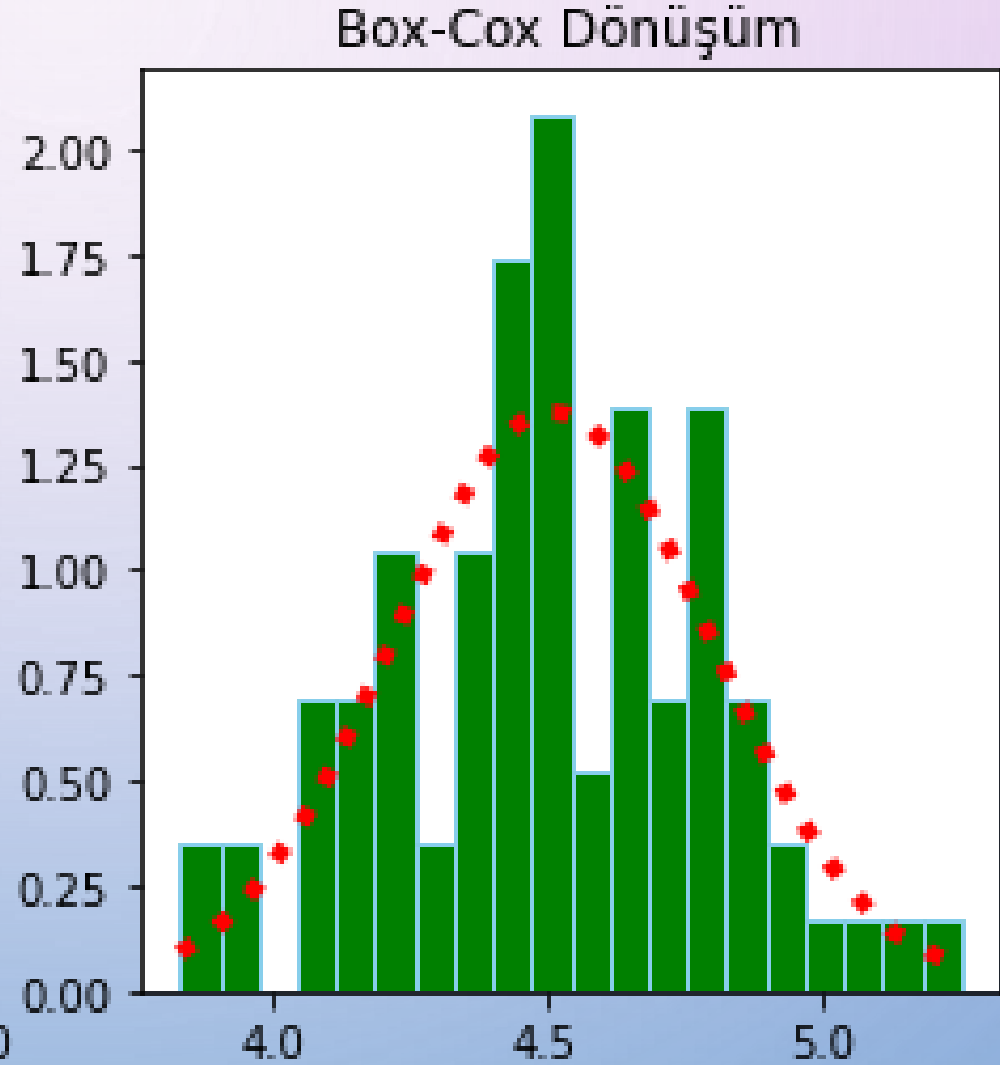
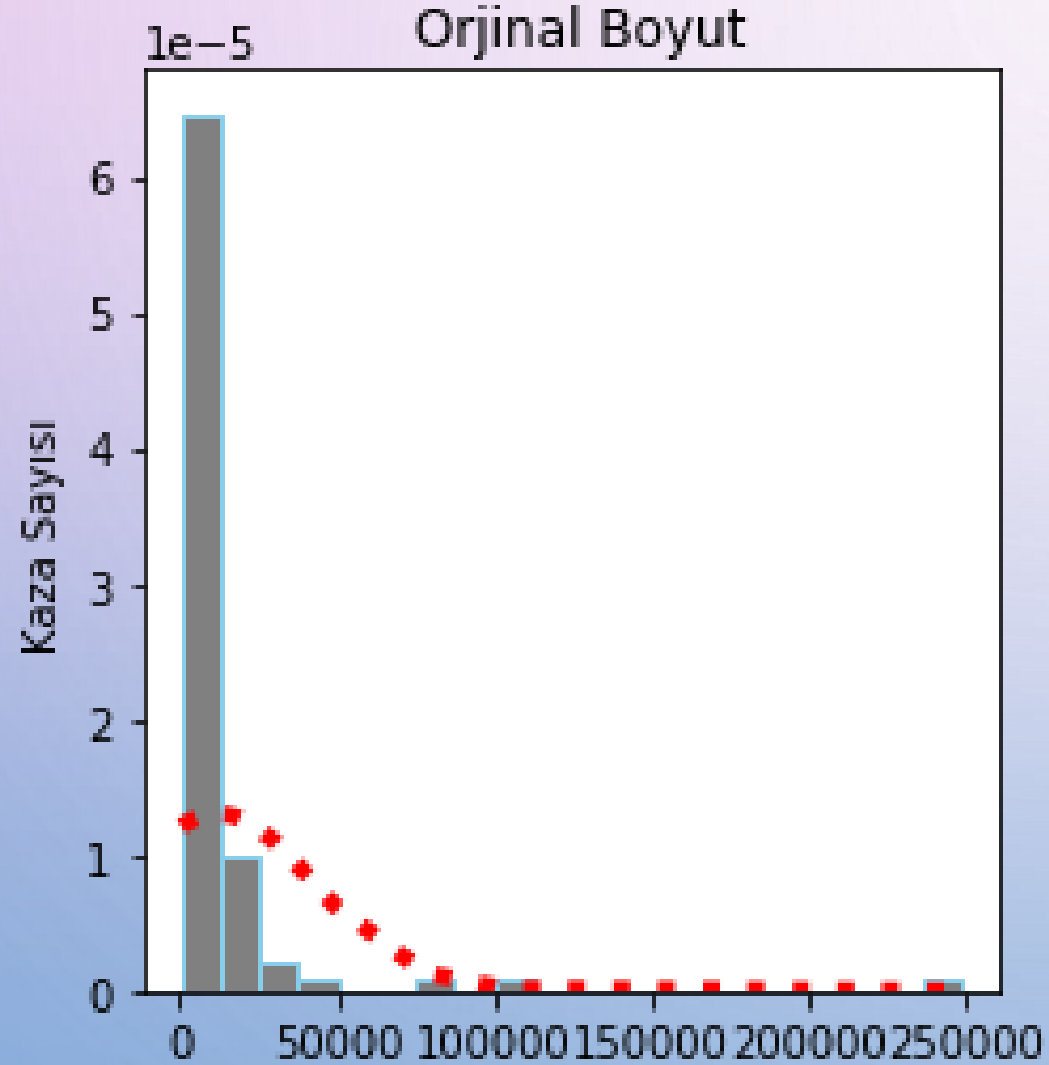
Verilerin birbirleri arasındaki ilişkilerini gösteren grafikler



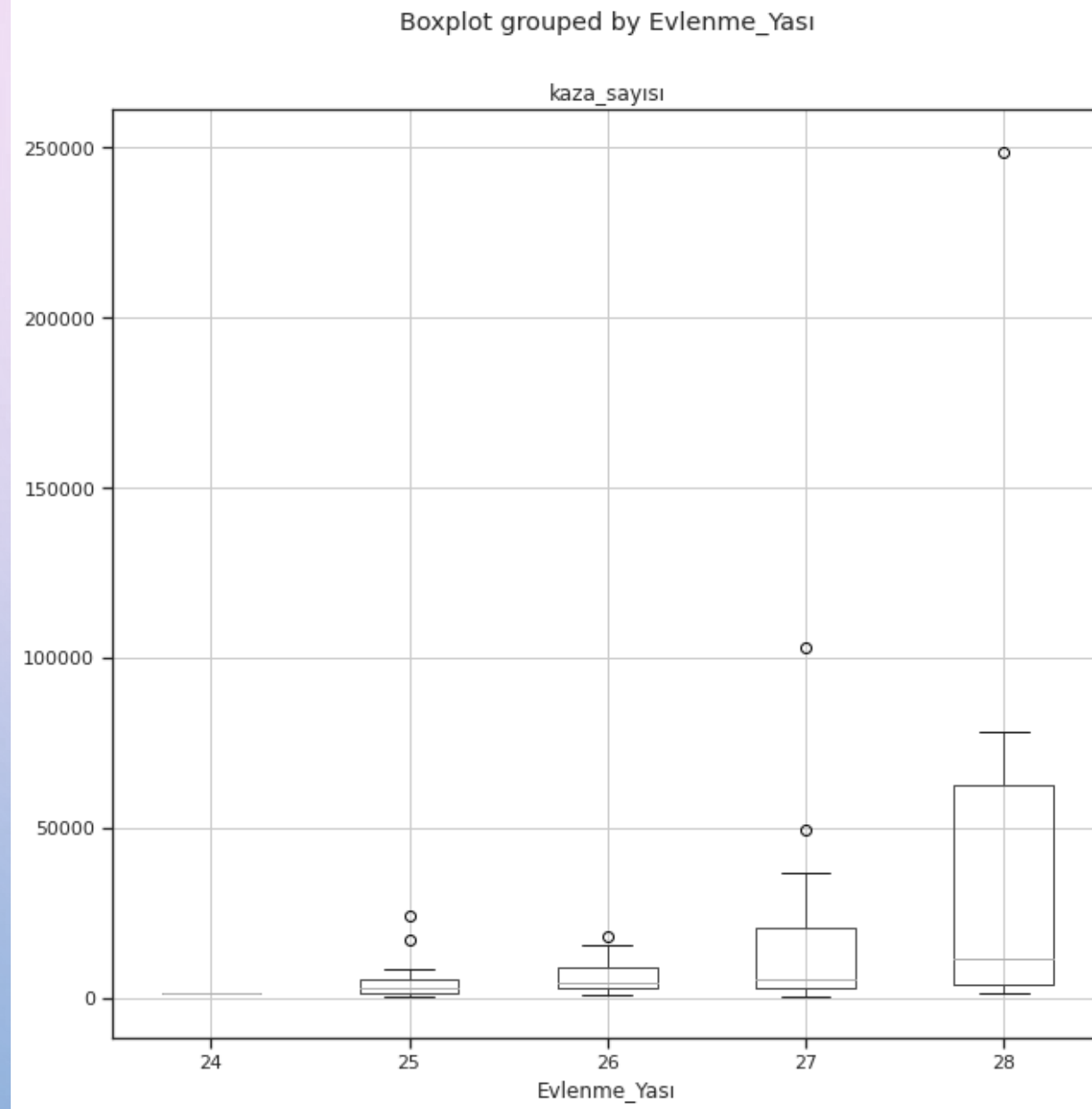
Boşanma sayısı ve kaza sayısı arasında %98 pozitif yönlü bir ilişki grafiği;



Eğer veriler normal dağılıma sahip olsaydı nasıl görünmesi gerektiğinin bir grafiği aşağıdaki gibidir.



Evlenme yaşı ile kaza sayısı arasındaki ilişkiyi gösteren grafikdir. Bu grafik sayesinde verinin içinde bulunan aykırı değerler gözlemlenebilir.

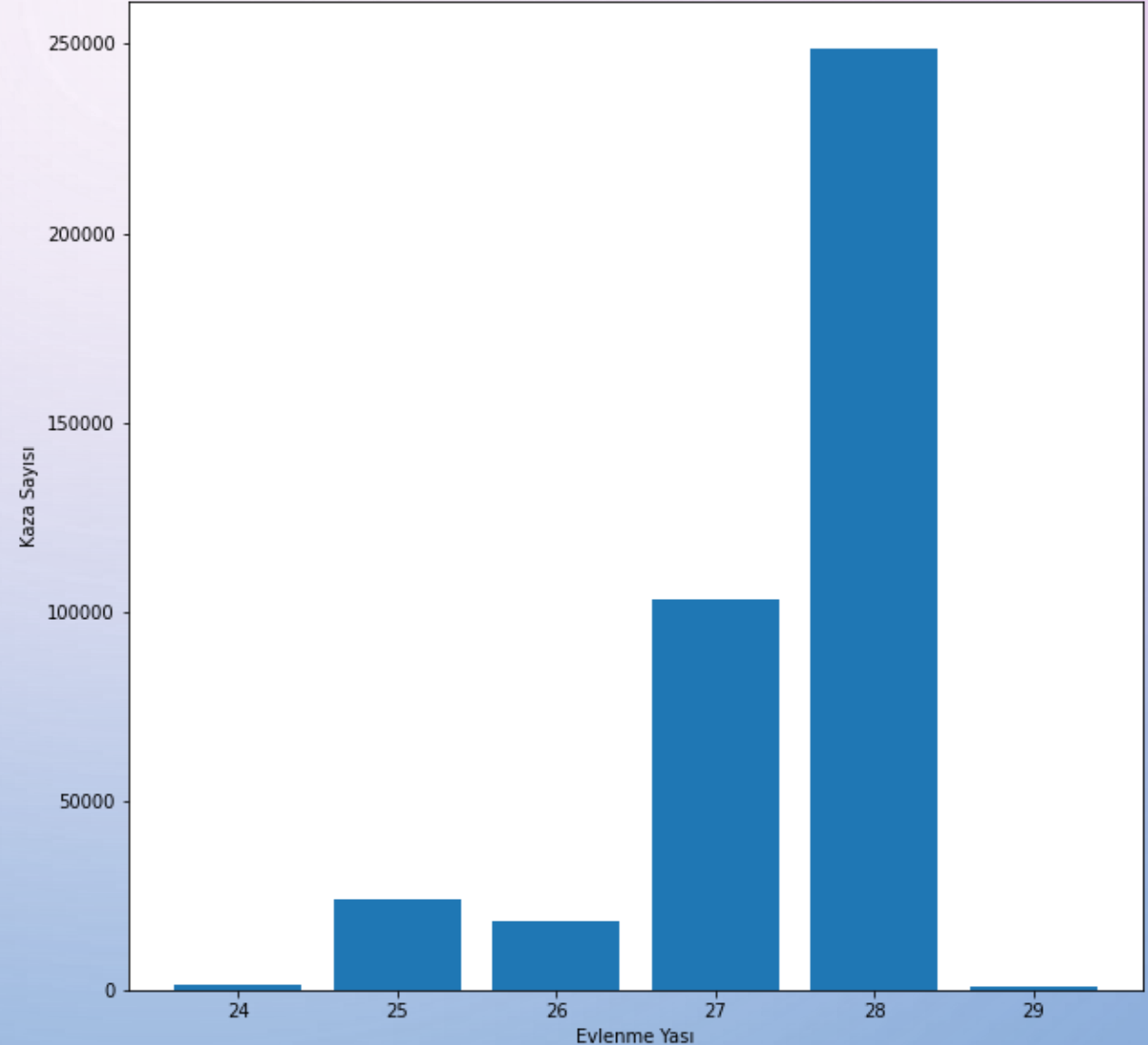


Yanda bulunan kutular verilerin varyanslarının genel olarak nerede olduğunu gösterir.

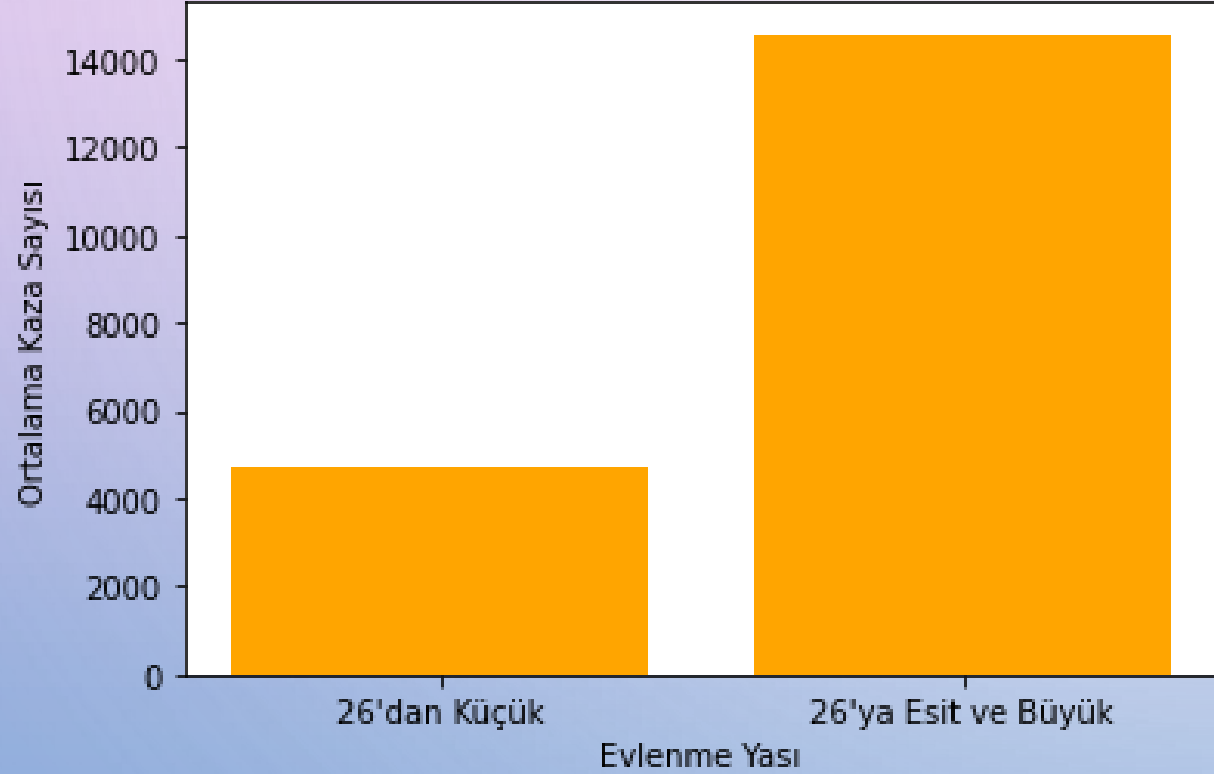
Elimizde bulunan veriler üstünde bazı örnekler üstünde değerlendirmeler yapacak olursak;

Evlenme yaşı ile kaza sayısı arasındaki ilişki;

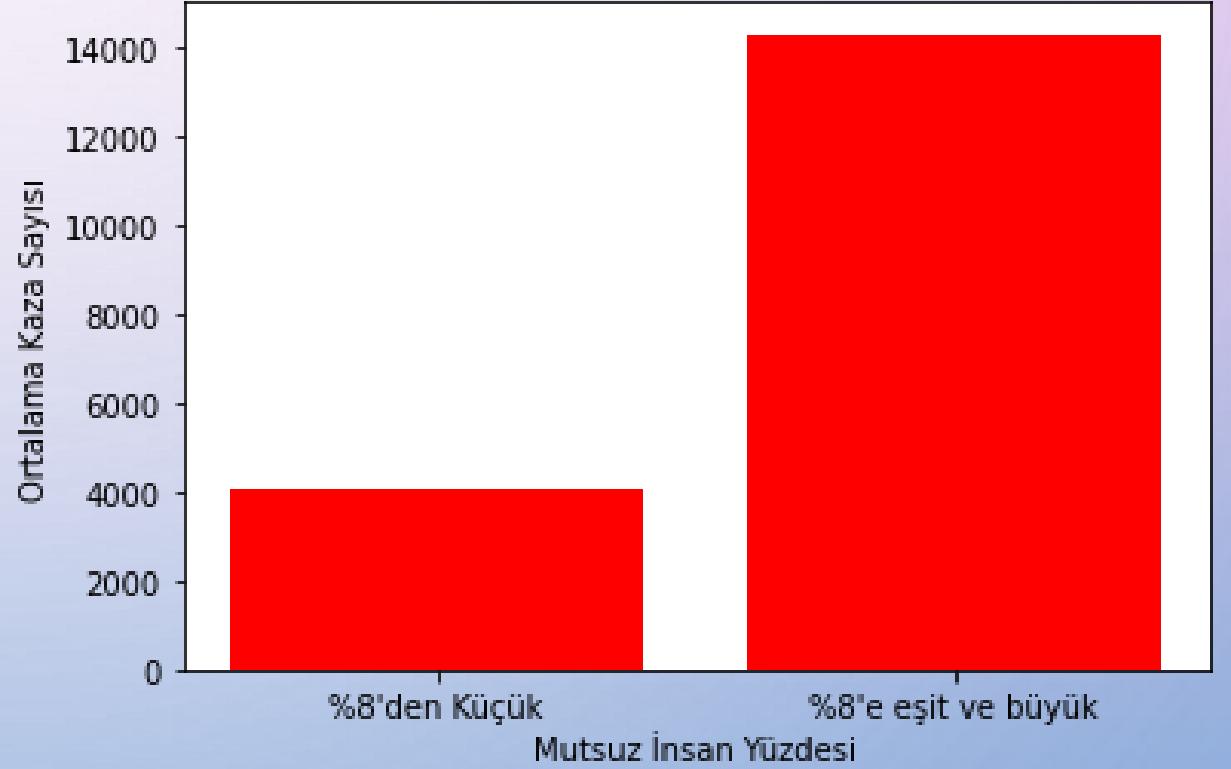
Burada görüldüğü üzere evlenme yaşı 28 olan insanların diğer insanlara göre kaza yapma olasılığı daha yüksektir.



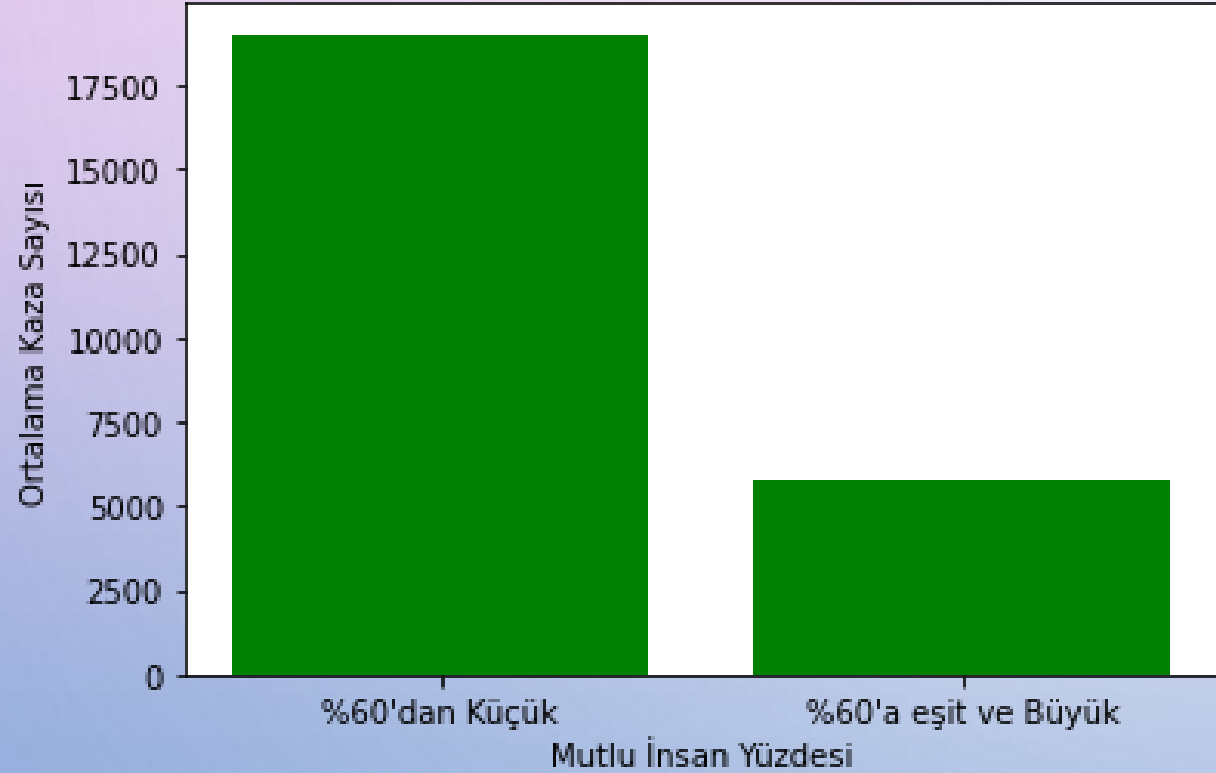
Ortalama kaza sayısı ile evlenme yaşı arasındaki ilişki



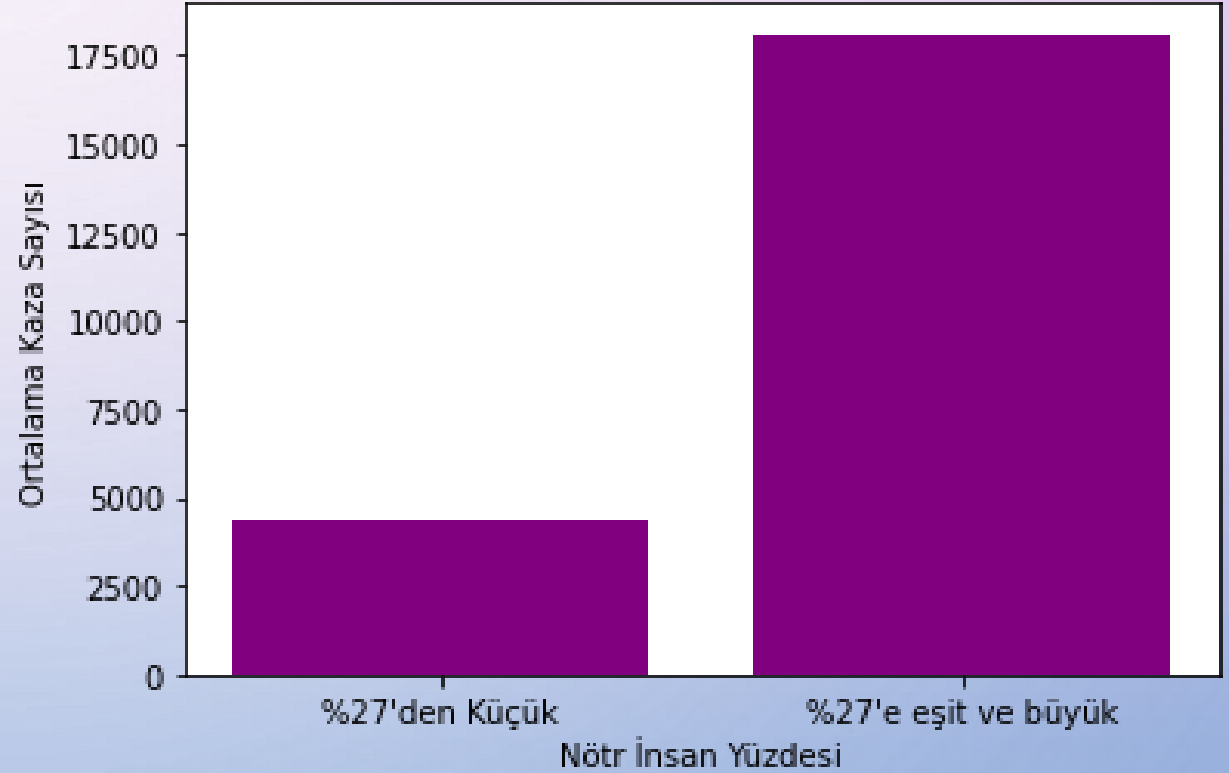
Ülke genelinde Mutsuz olan insanların kaza sayısının ortalaması ile arasındaki ilişki;



Mutlu olan insanların kaza sayısının ortalaması ile arasındaki ilişki;



Nötr olan insanların kaza sayısının ortalaması ile arasındaki ilişki;



Dinlediğiniz için teşekkür ederim.