
NOT - Bölüm-1’de detaylı inceleme yapmaya çalıştım. Veriyi üretme şeklimden yorumlara kadar en sağlıklı bölüm, Bölüm-1’dir.

NOT2 - Tüm ikili ve üçlü kombinasyonları ayrı başlıklar altında inceleyeceğim. Ancak kalabalık olmaması açısından kritik yerleri yorumlayacağım. Toplamda 10 tane çıktı inceleyeceğim.

NOT3 - Bölüm-2 den Bölüm-11’e kadar hep aynı işlemleri uyguladım. Bilgi kalabalığı olduğunu düşünürseniz, sol tarafta başlıklara ayırdığımdan dolayı en altta bulunan yorumlar kısmına geçiş yapabilirsiniz.

NOT4 - Başında “++” olan yeşil renkli yazılar benim çıkarımlarımdır.

BÖLÜM-1

++TC kimlik numaramın rakamları toplamını “yas” değişkenine atadım.
 ++Sonrasında rasgele 100 tane poisson dağılımına uygun sayıları, “yas” ortalaması ile üretip “hasar_sayisi” değişkenine atadım.

yas = (1+9+4+6+0+0+3+3+5+3+4)

hasar_sayisi = rpois(100,yas)

hasar_sayisi

[1] 40 40 42 46 40 44 40 42 40 36 32 41 43 38 39 44 41 40 31 32 47 42 42 41 42 36 24 34
 39 37 38 40

[33] 31 42 31 40 44 40 41 41 38 32 36 34 30 46 45 38 39 28 31 46 38 46 36 42 32 30 43 48
 35 30 32 51

[65] 39 34 23 39 35 38 45 44 36 37 33 34 45 43 41 37 53 35 41 37 40 33 32 38 39 33 35 44
 41 30 45 31

[97] 38 45 32 37

++Cinsiyet=gender, Araç Tipi = vehicle_type, Şehir = cities, Hasarsızlık = non_damaged isimleriyle rasgele sayılar ürettim. Bu sayıları seviyelere göre sırayla üretmemesi amacıyla, “sample” fonksiyonundan yardım aldım. İsim karışıklığı olmaması için tüm değişkenlerimin sonuna “2” rakamını ekledim. Rasgele üretilen verileri kendi içinde harmanlamış oldum.

>gender <- as.factor(rep(1:2, each=50, len=100))

>gender2 = sample(gender)

>gender2

[1] 1 1 1 1 2 1 1 1 1 2 2 1 1 2 1 2 2 2 1 1 1 2 1 2 2 2 1 1 2 2 1 1 2 2 2 1 1 2 2 1 1 2 2 2
 2

```
[49] 2 2 1 2 2 1 1 1 1 2 2 1 1 1 1 1 1 2 1 2 1 1 1 2 2 2 2 1 1 1 2 2 2 2 1 2 2 2 2 1 1 1 2
2 1
[97] 2 1 1 2
Levels: 1 2
```

```
> vehicle_type <- as.factor(rep(1:4, each=17, len=100))
> vehicle_type2 = sample(vehicle_type)
> vehicle_type2
[1] 4 2 2 3 2 2 2 2 1 1 2 2 3 2 2 1 2 4 2 2 4 1 2 1 4 1 2 2 2 1 2 1 2 4 3 2 3 3 1 1 3 2 3 4 4 3 4
4
[49] 4 3 2 1 3 2 2 1 1 2 1 2 2 4 3 1 1 3 1 1 4 1 4 3 1 1 1 1 4 1 1 1 2 2 1 4 3 1 2 1 3 1 1 3 1 4
2 1
[97] 1 2 3 4
Levels: 1 2 3 4
```

```
> cities <- as.factor(rep(1:5, each=6, len=100))
> cities2 = sample(cities)
> cities2
[1] 3 4 5 1 3 2 3 2 5 2 2 1 5 1 1 1 2 3 2 2 1 3 3 2 3 3 3 3 1 2 3 1 5 5 4 3 4 1 1 1 5 5 5 4 5 5 2
1
[49] 3 4 4 2 1 3 4 4 2 4 1 5 1 2 4 4 1 3 4 5 2 5 5 1 3 2 1 4 1 5 2 4 2 2 4 4 1 5 5 1 1 4 4 2 3 2
3 5
[97] 1 2 2 1
Levels: 1 2 3 4 5
```

```
non_damaged <- as.factor(rep(1:3, each=1, len=100))
> non_damaged2 = sample(non_damaged)
> non_damaged2
[1] 1 3 3 3 1 2 1 1 2 1 2 3 1 1 1 1 1 2 1 3 3 2 2 3 1 2 2 2 3 2 2 1 2 1 2 2 1 2 1 1 2 3 1 2 3 3 2
2
[49] 3 3 2 2 2 3 3 2 1 1 3 3 3 1 1 3 2 1 3 3 2 2 3 1 3 3 1 3 3 1 1 3 2 3 3 1 3 3 2 1 1 1 1 1 2 1
2 2
[97] 2 2 3 3
Levels: 1 2 3
```

++GLM hesabını yapmak için “glm” fonksiyonunu kullandım. Bu fonksiyonu da “proje” isminde kendi isimlendirdiğim değişkene atadım. GLM hesabında cinsiyet, araç tipi, hasarsızlık ve şehir verilerimin hepsini kullandım.

```
proje <- glm(hasar_sayisi ~ gender2 + vehicle_type2 + cities2 + non_damaged2 ,fam =
poisson(link = log))
```

++"summary()" komutu ile GLM değerlerini istemiş oldum.

>summary(proje)

Call:

```
glm(formula = hasar_sayisi ~ gender2 + vehicle_type2 + cities2 +  
    non_damaged2, family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.36364	-0.46835	-0.05003	0.55594	2.51098

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.7141807	0.0471473	78.778	<2e-16 ***
gender22	-0.0631188	0.0339200	-1.861	0.0628 .
vehicle_type22	-0.0231809	0.0413900	-0.560	0.5754
vehicle_type23	-0.0005625	0.0484786	-0.012	0.9907
vehicle_type24	0.0030717	0.0485905	0.063	0.9496
cities22	-0.0415947	0.0484917	-0.858	0.3910
cities23	-0.0212981	0.0527797	-0.404	0.6866
cities24	-0.1305948	0.0513368	-2.544	0.0110 *
cities25	-0.0631497	0.0509694	-1.239	0.2154
non_damaged22	0.0180751	0.0414536	0.436	0.6628
non_damaged23	0.0323114	0.0407960	0.792	0.4283

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 80.474 on 99 degrees of freedom
Residual deviance: 70.496 on 89 degrees of freedom
AIC: 640.22

++Deviance Residuals; artık değer anlamına gelir.

Çıktıda görüldüğü üzere;
minimum değerim; -2.36364,
1.çeyreklik; -0.46835
ortanca değer; -0.05003
3. çeyreklik; 0.55594
maksimum değerim; 2.51098

++Intercept; başlangıç değerimizdir. Çıktımızda ise başlangıç değerinin 3.7141807 olduğunu görüyoruz.

++GLM hesabı normal dağılıma göre yapılıyor ve normal dağılım negatif değer alabildiğinden dolayı (-) görebilmekteyiz.

++Signif. codes; hata payıdır. Hata payının düşük olması, gerçekleşen ve ortalama arasındaki farkı minimuma indirir. Sapma miktarı azalır.

++Değerin küçük olması beklenir. Ne kadar minimum ise hata payı da o kadar düşük olacaktır.

++Normal dağılım için $\alpha=0,05$ esas alınır ve modele alınacak değerin anlamlı olması için 0,05'ten küçük olması beklenir. Çıktımızda bu değeri sağlayan tek değişkenimiz var. İçlerinden "cities24" değişkeninin anlamlı olduğu sonucuna varabiliriz. Anlamlılığın az olması grafikte saçılımın fazla olması sonucuna ulaştırabilir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

++Residual Deviance ise hasar sayısının serbestlik derecesini verir.

Number of Fisher Scoring iterations: 4

> names(proje)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

++Katsayıları elde etmek için "coef" fonksiyonundan yardım alıyoruz.

> proje\$coef

```
(Intercept) gender22 vehicle_type22 vehicle_type23 vehicle_type24 cities22
3.7141807469 -0.0631187839 -0.0231808504 -0.0005625453 0.0030716545
-0.0415947372
cities23 cities24 cities25 non_damaged22 non_damaged23
-0.0212980878 -0.1305947628 -0.0631497427 0.0180750677 0.0323113805
```

$Y = 3.7141807469 + -0.0631187839(x_1) + -0.0231808504(x_2) + -0.0005625453(x_3) + \dots + 0.0323113805 (x_{10})$ denklemini kurabiliriz.

```
proje$fitted.values
```

[illegible]

```
proje$linear.predictors
```

1	2	3	4	5	6	7	8	9	10	11		
3.695954	3.592717	3.660162	3.745930	3.606583	3.667480	3.669702	3.649405	3.669106				
3.609467	3.604361											
12	13	14	15	16	17	18	19	20	21	22		
3.723311	3.650468	3.627881	3.691000	3.651062	3.586286	3.650911	3.649405	3.681717				
3.749564	3.647839											
23	24	25	26	27	28	29	30	31	32	33		
3.687777	3.641779	3.632836	3.647839	3.624658	3.687777	3.723311	3.627542	3.624658				
3.714181	3.582806											
34	35	36	37	38	39	40	41	42	43	44		
3.654103	3.537980	3.624658	3.583023	3.731693	3.714181	3.651062	3.605425	3.597043				
3.650468	3.604733											
45	46	47	48	49	50	51	52	53	54	55		
3.623295	3.619661	3.630614	3.672209	3.665147	3.552216	3.578480	3.627542	3.668574				
3.702013	3.592717											
56	57	58	59	60	61	62	63	64	65	66		

```

3.601661 3.672586 3.497286 3.683373 3.660162 3.723311 3.675658 3.583023 3.615897
3.732256 3.692320
  67   68   69   70   71   72   73   74   75   76   77
3.552779 3.683342 3.630614 3.669106 3.686414 3.713618 3.662075 3.641779 3.651062
3.552779 3.686445
  78   79   80   81   82   83   84   85   86   87   88
3.651031 3.672586 3.615897 3.604361 3.618598 3.552779 3.523539 3.682811 3.683342
3.582806 3.651062
  89   90   91   92   93   94   95   96   97   98   99
3.650499 3.520467 3.583586 3.672023 3.710958 3.612539 3.624658 3.669106 3.669137
3.667480 3.704335
 100
3.686445

```

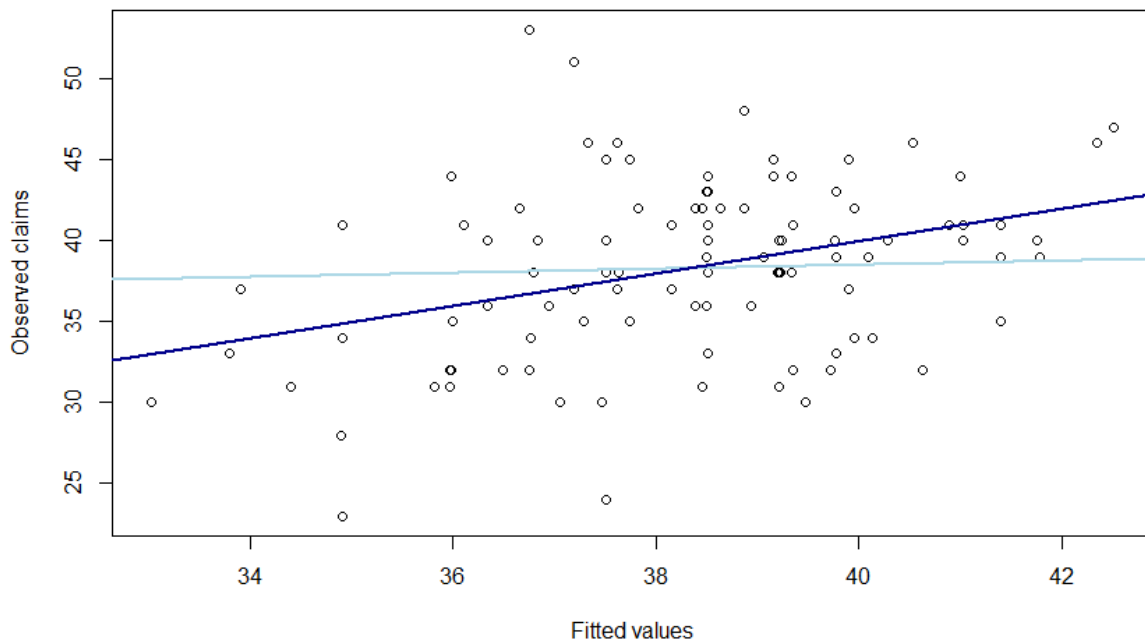
++Fit edilmiş değerler ile Lineer değerleri elde ettik. Lineer değerlerin diğer ismi ise öngörü değerleridir. Öngörü değerleri, kendi içinde daha tutarlı değerlere sahiptir.

++GLM hesabını grafikte görmek için "plot" komutunu kullandık. Doğrusallık hesabı için ise iki farklı çizgi üzerinden gözlem yapmak istedik. Açık maviye kadar olan kısım gözlemlerimizin sapma durumunu gösterecek. Lacivert ise doğrusallığın çıkarımını yaptıracak.

```

plot(proje$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
abline(lm(proje$fitted ~ hasar_sayisi), col="light blue", lwd=2)
abline(0, 1, col = "dark blue", lwd=2)

```



++Grafikten aldığım çıktıya göre; doğrusallığın olduğunu söyleyemiyorum. Saçılımlar çok fazla. Eğer daha doğrusal bir grafik elde etmek istiyorsam, $\alpha=0,05$ ten büyük olan değerleri verimden çıkarmam gerekirdi. Onun yerine ikili kombinasyonlar halinde analiz yapmaya devam edeceğim. Tekrar belirtmeliyim ki istediğim doğrusallık bu değil, şirket için bu veriler doğru sonuca ulaşmamda engel teşkil edecek.

++Son olarak "AIC" yardımıyla hesap yapacağız.

```
>AIC(proje)
[1] 640.2165
```

++640.2165 sonucunu aldık. Diğer değişkenlerle olan denememe göre hangi modeli seçeceğime karar vereceğim. AIC değerim ne kadar küçük olursa, o modeli seçmem daha doğru olacaktır.

```
anova(proje, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid.	Df	Resid.	Dev Pr(>Chi)
NULL				99	116.46	
gender2	1	0.0003		98	116.46	0.9871
vehicle_type2	3	2.7164		95	113.75	0.4374
cities2	4	3.3484		91	110.40	0.5013
non_damaged2	2	3.9588		89	106.44	0.1381

++Anova testini yaptık. $\alpha > 0,05$ olduğundan dolayı H_0 reddedilemez. Model anlamlıdır.

BÖLÜM-2

```
> proje1 <- glm(hasar_sayisi ~ gender2 + vehicle_type2 ,fam = poisson(link = log))
> summary(proje1)
```

Call:

```
glm(formula = hasar_sayisi ~ gender2 + vehicle_type2, family = poisson(link = log))
```

Deviance Residuals:

Min 1Q Median 3Q Max

-2.5290 -0.8930 0.1220 0.5927 3.0113

Coefficients:

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) 3.658115 0.034204 106.950 <2e-16 ***
gender22 -0.004673 0.032821 -0.142 0.887
vehicle_type22 -0.038567 0.040590 -0.950 0.342
vehicle_type23 0.039101 0.047232 0.828 0.408
vehicle_type24 0.002833 0.048180 0.059 0.953
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 116.46 on 99 degrees of freedom
Residual deviance: 113.75 on 95 degrees of freedom
AIC: 671.54

Number of Fisher Scoring iterations: 4

++Deviance Residuals; artık değer anlamına gelir.

Çıktıda görüldüğü üzere;

minimum değerim; -2.5290

1. çeyreklik; -0.8930

ortanca değer; 0.1220

3. çeyreklik; 0.5927

maksimum değerim; 3.0113

++Başlangıç değeri; 3.658115 tir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

> names(proje1)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```


> proje1\$coef

```
(Intercept)  gender22 vehicle_type22 vehicle_type23 vehicle_type24
3.658115163 -0.004673093 -0.038567407  0.039100681  0.002832899
```

++Y = 3.658115163 + -0.004673093(x1) + -0.038567407(x2) + 0.039100681(x3) + 0.002832899(x4) denklemini kurabilirsiniz.

> proje1\$fitted.values

```
  1    2    3    4    5    6    7    8    9   10   11
37.32069 37.32069 38.60733 37.32069 38.60733 37.32069 40.14680 38.60733 40.14680
37.14669 37.14669
 12   13   14   15   16   17   18   19   20   21   22
38.78816 38.60733 37.14669 37.14669 38.71685 38.78816 37.14669 40.33485 37.32069
40.33485 37.14669
 23   24   25   26   27   28   29   30   31   32   33
40.14680 37.32069 38.71685 38.60733 37.32069 40.14680 37.14669 38.78816 37.32069
37.14669 40.33485
 34   35   36   37   38   39   40   41   42   43   44
38.89820 38.89820 40.14680 38.89820 38.60733 38.89820 37.14669 38.60733 38.71685
38.60733 38.60733
 45   46   47   48   49   50   51   52   53   54   55
38.78816 37.32069 37.32069 40.33485 37.32069 38.71685 38.71685 40.14680 37.32069
37.32069 38.60733
 56   57   58   59   60   61   62   63   64   65   66
40.33485 38.78816 37.32069 37.14669 38.60733 37.32069 38.60733 38.78816 37.32069
38.60733 38.89820
 67   68   69   70   71   72   73   74   75   76   77
38.78816 38.89820 38.60733 38.78816 38.78816 40.14680 37.32069 38.78816 37.32069
40.33485 38.78816
 78   79   80   81   82   83   84   85   86   87   88
38.60733 38.89820 38.60733 37.14669 38.71685 38.89820 40.33485 38.89820 37.32069
38.60733 38.60733
 89   90   91   92   93   94   95   96   97   98   99
38.78816 38.78816 38.71685 38.60733 40.33485 38.60733 40.14680 38.60733 37.14669
37.14669 38.89820
100
```

40.14680

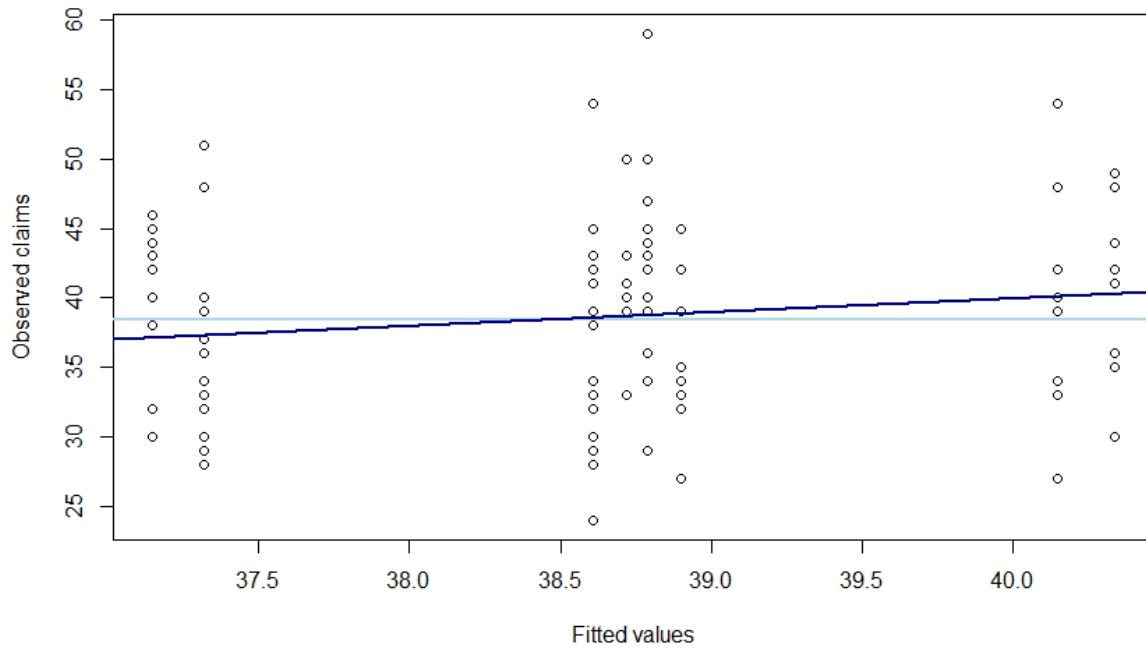
> proje1\$linear.predictors

1	2	3	4	5	6	7	8	9	10	11
3.619548	3.619548	3.653442	3.619548	3.653442	3.619548	3.692543	3.653442	3.692543	3.614875	3.614875
12	13	14	15	16	17	18	19	20	21	22
3.658115	3.653442	3.614875	3.614875	3.656275	3.658115	3.614875	3.697216	3.619548	3.697216	3.614875
23	24	25	26	27	28	29	30	31	32	33
3.692543	3.619548	3.656275	3.653442	3.619548	3.692543	3.614875	3.658115	3.619548	3.614875	3.697216
34	35	36	37	38	39	40	41	42	43	44
3.660948	3.660948	3.692543	3.660948	3.653442	3.660948	3.614875	3.653442	3.656275	3.653442	3.653442
45	46	47	48	49	50	51	52	53	54	55
3.658115	3.619548	3.619548	3.697216	3.619548	3.656275	3.656275	3.692543	3.619548	3.619548	3.653442
56	57	58	59	60	61	62	63	64	65	66
3.697216	3.658115	3.619548	3.614875	3.653442	3.619548	3.653442	3.658115	3.619548	3.653442	3.660948
67	68	69	70	71	72	73	74	75	76	77
3.658115	3.660948	3.653442	3.658115	3.658115	3.692543	3.619548	3.658115	3.619548	3.697216	3.658115
78	79	80	81	82	83	84	85	86	87	88
3.653442	3.660948	3.653442	3.614875	3.656275	3.660948	3.697216	3.660948	3.619548	3.653442	3.653442
89	90	91	92	93	94	95	96	97	98	99
3.658115	3.658115	3.656275	3.653442	3.697216	3.653442	3.692543	3.653442	3.614875	3.614875	3.660948
100										
3.692543										

> plot(proje1\$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")

> abline(lm(proje1\$fitted ~ hasar_sayisi), col="light blue", lwd=2)

> abline(0, 1, col = "dark blue", lwd=2)



++Plot yardımıyla çizdiğimiz grafik, yine doğrusallığın olmadığını gösteriyor. Veriler 3 gruba ayrılmış bir halde saçılmış. İstedğimiz görüntü, lacivert ve mavi çizginin eğimi etrafında yayılan daireler görmektir.

```
> AIC(proje1)
[1] 671.5448
```

++AIC değeri ilk testime göre daha yüksek çıktı. Bizim isteğimiz ise daha düşük olmasıydı. Bu nedenle bu deneyi kullanmak doğru olmayacaktır.

```
> anova(proje1, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			99	116.46	
gender2	1	0.00026	98	116.46	0.9871
vehicle_type2	3	2.71639	95	113.75	0.4374

++Alfa>0,05 sonucunu aldık. Ho reddedilemez. Model anlamlıdır.

BÖLÜM-3

```
> proje2 <- glm(hasar_sayisi ~ non_damaged2 + vehicle_type2 ,fam = poisson(link = log))
> summary(proje2)
```

Call:

```
glm(formula = hasar_sayisi ~ non_damaged2 + vehicle_type2, family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.3682	-0.7935	0.1646	0.6692	2.8264

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.685555	0.036350	101.391	<2e-16 ***
non_damaged22	-0.060097	0.039611	-1.517	0.129
non_damaged23	-0.026651	0.039458	-0.675	0.499
vehicle_type22	-0.039411	0.040016	-0.985	0.325
vehicle_type23	0.032465	0.047370	0.685	0.493
vehicle_type24	0.002549	0.048110	0.053	0.958

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 116.46 on 99 degrees of freedom
Residual deviance: 111.46 on 94 degrees of freedom
AIC: 671.26

Number of Fisher Scoring iterations: 4

++Deviance Residuals; artık değer anlamına gelir.
Çıktıda görüldüğü üzere;

minimum değerim; -2.3682
1. çeyreklik; -0.7935
ortanca değer; 0.1646
3. çeyreklik; 0.6692
maksimum değerim; 2.8264

++Başlangıç değeri; 3.685555 tir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

```
> names(proje2)
```

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

```
> proje2$coef
```

```
(Intercept) non_damaged22 non_damaged23 vehicle_type22 vehicle_type23
vehicle_type24
3.685554677 -0.060097435 -0.026651121 -0.039411252 0.032464760
0.002548588
```

$++Y = 3.685554677 + -0.060097435(x1) + -0.026651121(x2) + -0.039411252(x3) + 0.032464760(x4)$ denklemini kurabiliriz.

```
> proje2$fitted.values
```

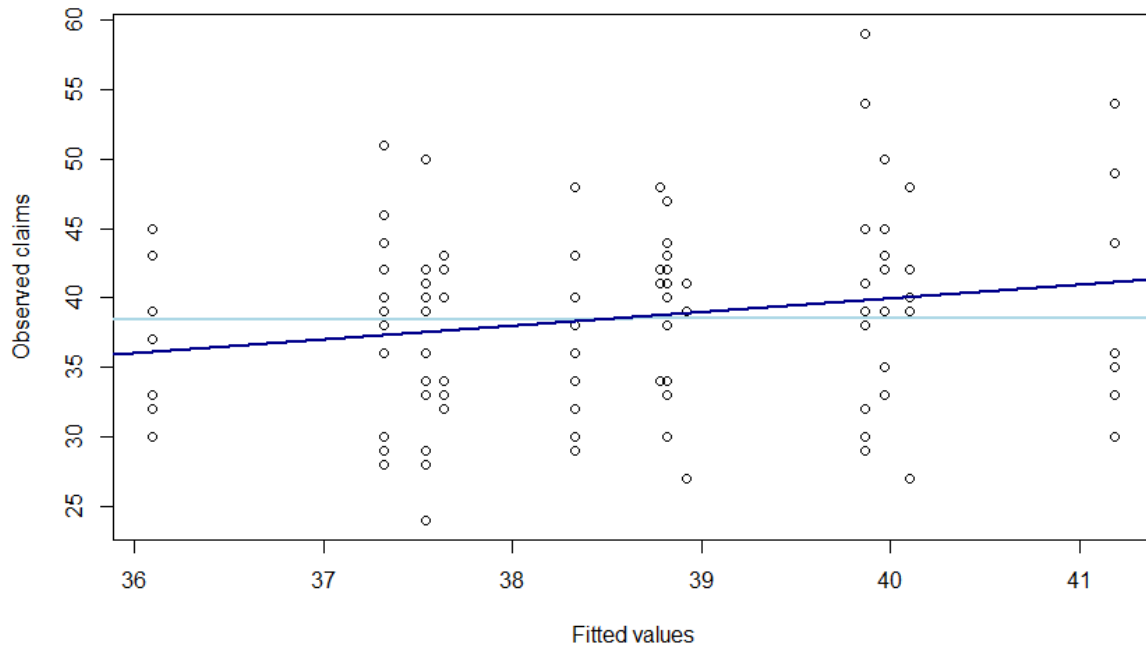
```
1 2 3 4 5 6 7 8 9 10 11
38.32657 37.31862 37.54188 36.09109 37.54188 38.32657 40.09968 37.54188 40.09968
37.31862 38.32657
12 13 14 15 16 17 18 19 20 21 22
38.81876 39.86723 38.32657 38.32657 38.91782 38.81876 37.31862 41.18275 38.32657
41.18275 37.31862
23 24 25 26 27 28 29 30 31 32 33
40.09968 36.09109 37.63769 39.86723 36.09109 41.18275 37.31862 38.81876 37.31862
37.31862 41.18275
34 35 36 37 38 39 40 41 42 43 44
39.96896 39.96896 40.09968 37.63769 37.54188 38.91782 38.32657 38.81876 37.63769
38.81876 38.81876
45 46 47 48 49 50 51 52 53 54 55
38.81876 36.09109 37.31862 41.18275 37.31862 39.96896 38.91782 38.78067 36.09109
38.32657 37.54188
56 57 58 59 60 61 62 63 64 65 66
40.09968 39.86723 36.09109 37.31862 38.81876 37.31862 38.81876 37.54188 36.09109
37.54188 37.63769
67 68 69 70 71 72 73 74 75 76 77
39.86723 39.96896 39.86723 39.86723 37.54188 38.78067 36.09109 39.86723 37.31862
41.18275 37.54188
78 79 80 81 82 83 84 85 86 87 88
37.54188 37.63769 38.81876 38.32657 39.96896 37.63769 38.78067 39.96896 38.32657
39.86723 37.54188
```

```
      89      90      91      92      93      94      95      96      97      98      99
38.81876 37.54188 39.96896 38.81876 40.09968 39.86723 41.18275 39.86723 36.09109
36.09109 37.63769
      100
38.78067
```

```
> proje2$linear.predictors
```

```
      1      2      3      4      5      6      7      8      9     10     11
3.646143 3.619492 3.625457 3.586046 3.625457 3.646143 3.691368 3.625457 3.691368
3.619492 3.646143
      12     13     14     15     16     17     18     19     20     21     22
3.658904 3.685555 3.646143 3.646143 3.661452 3.658904 3.619492 3.718019 3.646143
3.718019 3.619492
      23     24     25     26     27     28     29     30     31     32     33
3.691368 3.586046 3.628006 3.685555 3.586046 3.718019 3.619492 3.658904 3.619492
3.619492 3.718019
      34     35     36     37     38     39     40     41     42     43     44
3.688103 3.688103 3.691368 3.628006 3.625457 3.661452 3.646143 3.658904 3.628006
3.658904 3.658904
      45     46     47     48     49     50     51     52     53     54     55
3.658904 3.586046 3.619492 3.718019 3.619492 3.688103 3.661452 3.657922 3.586046
3.646143 3.625457
      56     57     58     59     60     61     62     63     64     65     66
3.691368 3.685555 3.586046 3.619492 3.658904 3.619492 3.658904 3.625457 3.586046
3.625457 3.628006
      67     68     69     70     71     72     73     74     75     76     77
3.685555 3.688103 3.685555 3.685555 3.625457 3.657922 3.586046 3.685555 3.619492
3.718019 3.625457
      78     79     80     81     82     83     84     85     86     87     88
3.625457 3.628006 3.658904 3.646143 3.688103 3.628006 3.657922 3.688103 3.646143
3.685555 3.625457
      89     90     91     92     93     94     95     96     97     98     99
3.658904 3.625457 3.688103 3.658904 3.691368 3.685555 3.718019 3.685555 3.586046
3.586046 3.628006
      100
3.657922
```

```
> plot(proje2$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje2$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Bu sefer daha dağınık bir grafikte karşı karşıyayız. Henüz istediğimiz doğrusallığı yakalayamadık. İki değişken arası bağ bulunmuyor gibi gözükmemekte.

```
> AIC(proje2)
[1] 671.2567
```

++AIC değerim bir önceki test ile aynı sayılır. Hala ilk modelimi tercih etmem gerektiğini görüyorum.

```
> anova(proje2, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

		Df	Deviance	Resid.	Df	Resid.	Dev	Pr(>Chi)
NULL		99		116.46				
non_damaged2	2	2.5759	97	113.89	0.2758			
vehicle_type2	3	2.4289	94	111.46	0.4883			

++Alfa>0,05, model anlamlıdır. Ho reddedilemez.

BÖLÜM-4

```
> proje3 <- glm(hasar_sayisi ~ non_damaged2 + gender2 ,fam = poisson(link = log))
> summary(proje3)
```

Call:

```
glm(formula = hasar_sayisi ~ non_damaged2 + gender2, family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.45214	-0.70760	-0.05422	0.80028	2.03128

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.59991	0.03262	110.372	<2e-16 ***
non_damaged22	0.04640	0.03939	1.178	0.2388
non_damaged23	0.01041	0.03988	0.261	0.7941
gender22	0.06197	0.03275	1.892	0.0585 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.21 on 99 degrees of freedom
 Residual deviance: 102.38 on 96 degrees of freedom
 AIC: 658.22

Number of Fisher Scoring iterations: 4

++Deviance Residuals; artık değer anlamına gelir.
 Çıktıda görüldüğü üzere;

minimum değerim;-2.45214
 1.çeyreklik;-0.70760
 ortanca değer; -0.05422
 3. çeyreklik; 0.80028
 maksimum değerim; 2.03128

++Başlangıç değeri; 3.59991 tir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

> names(proje3)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

> proje3\$coef

```
(Intercept) non_damaged22 non_damaged23 gender22
3.59991079 0.04639986 0.01040798 0.06196847
```

++ $Y = 3.59991079 + 0.04639986(x_1) + 0.01040798(x_2) + 0.06196847(x_3)$

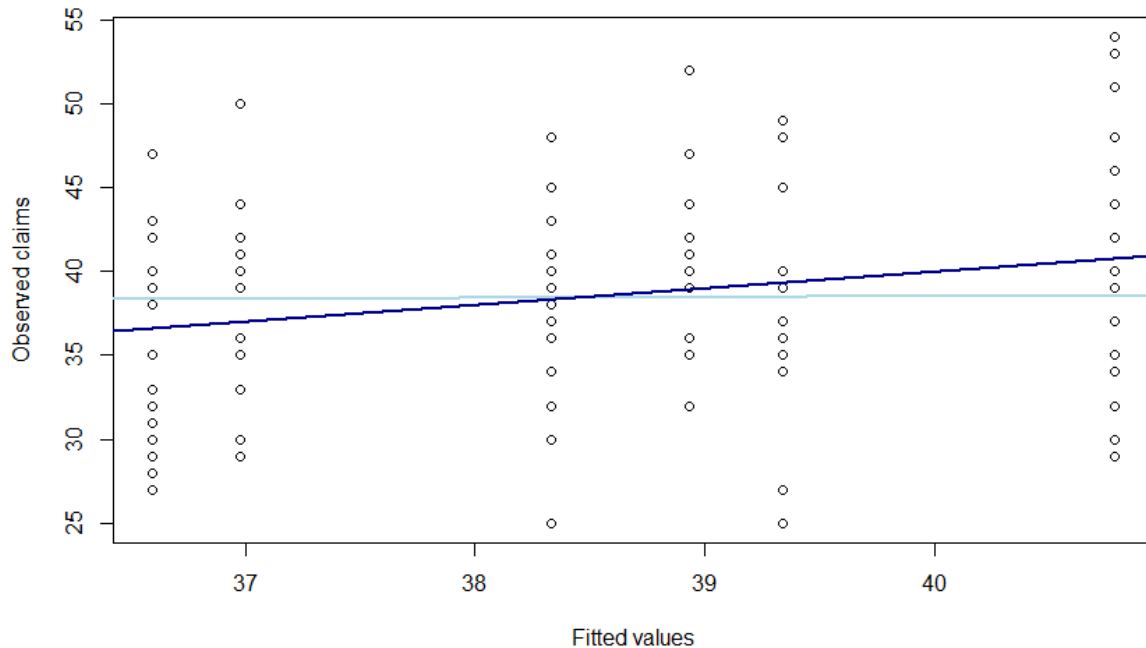
denklemini kurabiliriz.

> proje3\$fitted.values

```
 1  2  3  4  5  6  7  8  9 10 11
38.33298 39.34179 36.59497 40.78356 38.93444 39.34179 38.33298 40.78356 38.93444
38.93444 36.59497
12 13 14 15 16 17 18 19 20 21 22
36.97784 38.93444 38.93444 36.59497 36.59497 39.34179 38.93444 38.33298 36.97784
36.97784 36.59497
23 24 25 26 27 28 29 30 31 32 33
40.78356 36.59497 36.97784 36.97784 39.34179 36.59497 38.33298 36.97784 39.34179
40.78356 36.97784
34 35 36 37 38 39 40 41 42 43 44
38.33298 38.33298 36.59497 38.93444 36.97784 40.78356 40.78356 36.97784 36.59497
36.97784 39.34179
45 46 47 48 49 50 51 52 53 54 55
39.34179 40.78356 36.97784 39.34179 38.93444 38.93444 38.93444 36.59497 40.78356
39.34179 36.97784
56 57 58 59 60 61 62 63 64 65 66
38.33298 40.78356 38.93444 40.78356 38.33298 40.78356 38.33298 40.78356 36.97784
39.34179 36.59497
67 68 69 70 71 72 73 74 75 76 77
36.97784 38.93444 39.34179 38.93444 40.78356 38.33298 36.97784 36.97784 38.93444
40.78356 38.33298
```



```
> plot(proje3$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje3$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Grafikimiz yine istediğimiz şekli vermedi. Doğrusallık sağlanmadı.

```
> AIC(proje3)
[1] 658.2231
```

++AIC değerimiz bu sefer biraz daha düşük geldi. Ancak ilk örneklem hala tercih etmek için iyi bir konumda.

```
> anova(proje3, test="Chisq")
```

Analysis of Deviance Table

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			99	108.21	
non_damaged2	2	2.2538	97	105.96	0.3240

```
gender2    1  3.5823    96  102.38  0.0584 .
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++Alfa>0.05 , Ho reddedilemez. Model anlamlıdır.

BÖLÜM-5

```
> proje4 <- glm(hasar_sayisi ~ non_damaged2 + cities2 ,fam = poisson(link = log))
> summary(proje4)
```

Call:

```
glm(formula = hasar_sayisi ~ non_damaged2 + cities2, family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.17098	-0.69466	0.05551	0.66107	2.09240

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.684251	0.040951	89.968	<2e-16 ***
non_damaged22	0.049545	0.040269	1.230	0.2186
non_damaged23	-0.005623	0.040334	-0.139	0.8891
cities22	-0.109999	0.048890	-2.250	0.0245 *
cities23	-0.016447	0.050933	-0.323	0.7468
cities24	-0.012021	0.049962	-0.241	0.8099
cities25	-0.114438	0.051053	-2.242	0.0250 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.214 on 99 degrees of freedom
Residual deviance: 96.154 on 93 degrees of freedom
AIC: 658

Number of Fisher Scoring iterations: 4

++Deviance Residuals; artık değer anlamına gelir.
Çıktıda görüldüğü üzere;

minimum değerim; -2.17098
1.çeyreklik; -0.69466
ortanca değer; 0.05551
3. çeyreklik; 0.66107

maksimum değerim; 2.09240

++Başlangıç değeri;3.684251 tir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

> names(proje4)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

> proje4\$coef

```
(Intercept) non_damaged22 non_damaged23 cities22 cities23 cities24 cities25
3.684251276 0.049545430 -0.005622864 -0.109998974 -0.016446793 -0.012020808
-0.114438373
```

++ $Y = 3.684251276 + 0.049545430(x_1) + -0.005622864(x_2) + -0.109998974(x_3) + -0.016446793(x_4) + -0.012020808(x_5) + -0.114438373(x_6)$
denklemini kurabiliriz.

> proje4\$fitted.values

```
1 2 3 4 5 6 7 8 9 10 11
41.83765 39.59205 35.50995 41.15518 39.81530 35.31084 41.15518 41.15518 35.50995
35.50995 35.66794
12 13 14 15 16 17 18 19 20 21 22
39.59205 39.16582 39.33955 39.33955 35.66794 35.46795 39.81530 37.31362 35.31084
39.59205 35.66794
23 24 25 26 27 28 29 30 31 32 33
37.47964 39.33955 39.59205 38.94622 35.46795 35.66794 41.15518 39.59205 35.31084
37.31362 39.59205
```

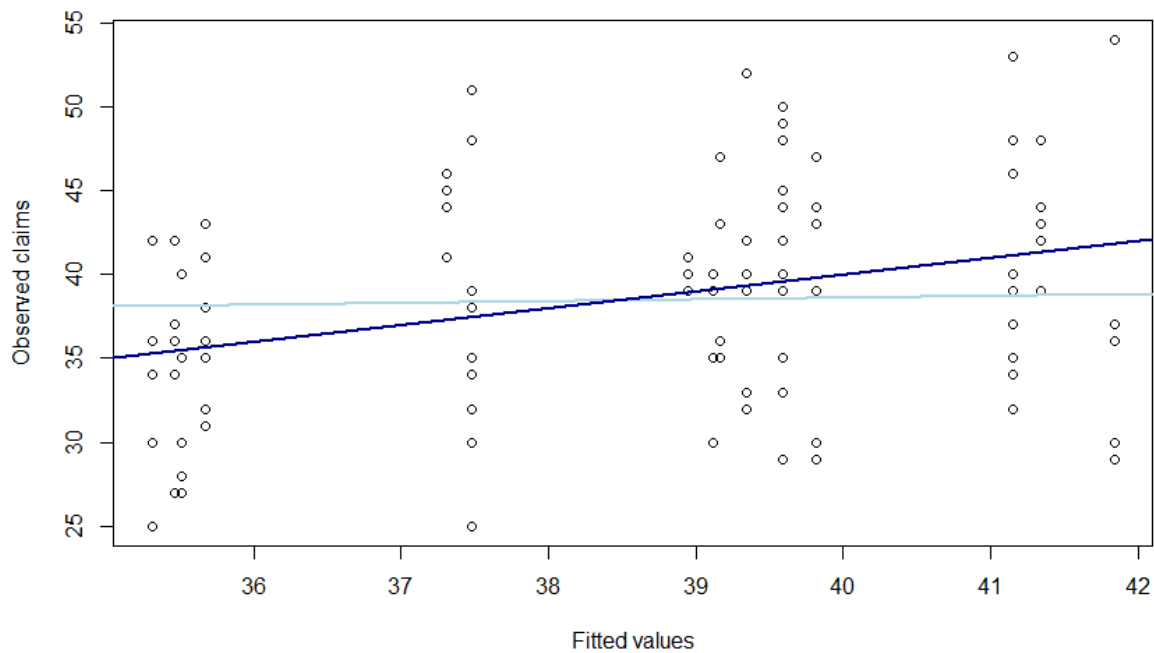
34	35	36	37	38	39	40	41	42	43	44
41.83765	37.47964	35.66794	35.66794	39.11897	41.15518	37.31362	35.31084	39.33955	39.11897	39.59205
45	46	47	48	49	50	51	52	53	54	55
35.46795	41.15518	39.59205	35.31084	39.33955	35.50995	35.66794	39.33955	41.33774	38.94622	39.11897
56	57	58	59	60	61	62	63	64	65	66
37.47964	37.47964	35.66794	41.83765	41.33774	41.15518	37.47964	41.15518	35.46795	39.11897	39.81530
67	68	69	70	71	72	73	74	75	76	77
35.46795	35.50995	39.59205	39.33955	41.33774	37.31362	39.59205	38.94622	39.33955	37.47964	37.47964
78	79	80	81	82	83	84	85	86	87	88
38.94622	39.59205	39.81530	39.11897	39.81530	41.83765	39.16582	41.15518	35.50995	41.33774	35.50995
89	90	91	92	93	94	95	96	97	98	99
39.81530	39.16582	41.15518	39.16582	41.33774	35.31084	39.81530	35.31084	37.47964	37.47964	41.83765
100										
39.59205										

> proje4\$linear.predictors

1	2	3	4	5	6	7	8	9	10	11
3.733797	3.678628	3.569813	3.717350	3.684251	3.564190	3.717350	3.717350	3.569813	3.569813	3.574252
12	13	14	15	16	17	18	19	20	21	22
3.678628	3.667804	3.672230	3.672230	3.574252	3.568629	3.684251	3.619358	3.564190	3.678628	3.574252
23	24	25	26	27	28	29	30	31	32	33
3.623798	3.672230	3.678628	3.662182	3.568629	3.574252	3.717350	3.678628	3.564190	3.619358	3.678628
34	35	36	37	38	39	40	41	42	43	44
3.733797	3.623798	3.574252	3.574252	3.666608	3.717350	3.619358	3.564190	3.672230	3.666608	3.678628
45	46	47	48	49	50	51	52	53	54	55
3.568629	3.717350	3.678628	3.564190	3.672230	3.569813	3.574252	3.672230	3.721776	3.662182	3.666608
56	57	58	59	60	61	62	63	64	65	66
3.623798	3.623798	3.574252	3.733797	3.721776	3.717350	3.623798	3.717350	3.568629	3.666608	3.684251
67	68	69	70	71	72	73	74	75	76	77
3.568629	3.569813	3.678628	3.672230	3.721776	3.619358	3.678628	3.662182	3.672230	3.623798	3.623798
78	79	80	81	82	83	84	85	86	87	88
3.662182	3.678628	3.684251	3.666608	3.684251	3.733797	3.667804	3.717350	3.569813	3.721776	3.569813
89	90	91	92	93	94	95	96	97	98	99

```
3.684251 3.667804 3.717350 3.667804 3.721776 3.564190 3.684251 3.564190 3.623798
3.623798 3.733797
100
3.678628
```

```
> plot(proje4$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje4$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Grafğimiz doğrusallığa uymamaktadır.

```
> AIC(proje4)
[1] 657.9999
```

++AIC değerimiz yeterince küçük gelmedi.

```
> anova(proje4, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

		Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL		99	108.214			
non_damaged2	2	2.2538	97	105.960	0.32404	
cities2	4	9.8055	93	96.154	0.04383	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++cities2 değişkeninde $\alpha < 0,05$. H_0 reddedilir. Model anlamlı değildir.

BÖLÜM-6

```
> proje5 <- glm(hasar_sayisi ~ vehicle_type2 + cities2 ,fam = poisson(link = log))
> summary(proje5)
```

Call:

```
glm(formula = hasar_sayisi ~ vehicle_type2 + cities2, family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.0931	-0.6048	-0.1250	0.6488	2.0647

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.732346	0.039385	94.766	<2e-16 ***
vehicle_type22	-0.016422	0.040975	-0.401	0.6886
vehicle_type23	-0.056536	0.048065	-1.176	0.2395
vehicle_type24	-0.122499	0.049257	-2.487	0.0129 *
cities22	-0.104270	0.048030	-2.171	0.0299 *
cities23	-0.007348	0.050037	-0.147	0.8833
cities24	-0.015122	0.049487	-0.306	0.7599
cities25	-0.119427	0.051711	-2.310	0.0209 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.214 on 99 degrees of freedom
 Residual deviance: 91.295 on 92 degrees of freedom
 AIC: 655.14

Number of Fisher Scoring iterations: 4

minimum değerim; -2.0931
 1.çeyreklik; -0.6048
 ortanca değer; -0.1250
 3. çeyreklik; 0.6488
 maksimum değerim; 2.0647

++Başlangıç değeri; 3.732346 tir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

> names(proje5)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

> proje5\$coef

```
(Intercept) vehicle_type22 vehicle_type23 vehicle_type24 cities22 cities23
3.732345691 -0.016422256 -0.056535657 -0.122499452 -0.104269860
-0.007347588
cities24 cities25
-0.015122144 -0.119427147
```

++ $Y = 3.732345691 + -0.016422256 (x_1) + -0.056535657 (x_2) + -0.122499452 (x_3) + -0.104269860 (x_4) + -0.007347588 (x_5) + -0.015122144 (x_6) + -0.119427147 (x_7)$

denklemini kurduk.

> proje5\$fitted.values

```
1 2 3 4 5 6 7 8 9 10 11
39.48062 39.48062 36.47023 40.79567 39.48062 36.47023 40.79567 40.79567 35.03624
36.47023 37.64032
12 13 14 15 16 17 18 19 20 21 22
39.48062 39.19160 40.47973 36.40565 37.64032 37.64032 41.77699 36.47023 36.47023
41.77699 35.57134
23 24 25 26 27 28 29 30 31 32 33
37.02723 41.14998 41.09652 41.47115 33.30063 35.57134 36.68980 41.77699 35.03624
36.47023 41.77699
34 35 36 37 38 39 40 41 42 43 44
41.09652 33.30063 35.57134 37.02723 38.88808 40.79567 37.07410 32.79969 41.14998
38.88808 39.48062
45 46 47 48 49 50 51 52 53 54 55
```

```

37.64032 40.79567 41.77699 37.07410 38.88808 37.07410 37.64032 40.47973 36.40565
40.79567 41.14998
  56    57    58    59    60    61    62    63    64    65    66
37.64032 37.64032 37.64032 36.96037 40.47973 39.19160 37.02723 41.47115 33.30063
41.14998 41.09652
  67    68    69    70    71    72    73    74    75    76    77
33.30063 36.47023 41.77699 41.14998 40.47973 36.47023 36.96037 41.47115 36.40565
37.02723 35.57134
  78    79    80    81    82    83    84    85    86    87    88
40.79567 41.09652 41.09652 41.14998 36.96037 36.96037 41.47115 40.79567 32.79969
41.14998 32.79969
  89    90    91    92    93    94    95    96    97    98    99
41.77699 36.68980 39.19160 40.79567 41.14998 36.47023 41.09652 37.07410 37.64032
37.64032 41.77699
  100
36.96037

```

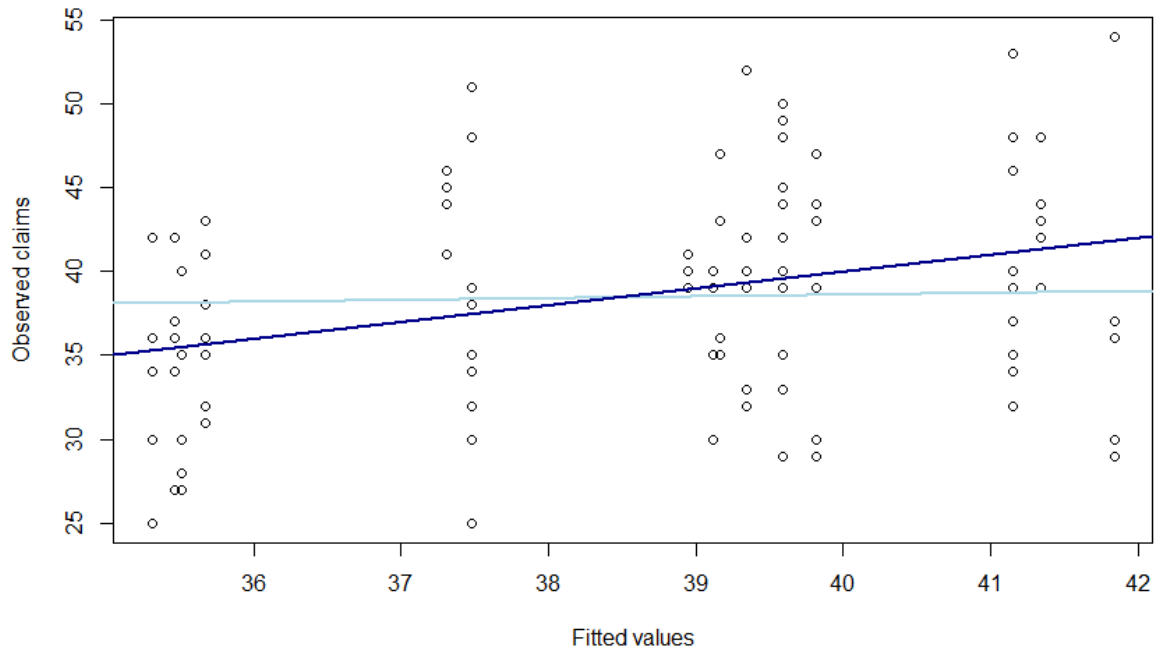
> proje5\$linear.predictors

```

  1    2    3    4    5    6    7    8    9   10   11
3.675810 3.675810 3.596496 3.708576 3.675810 3.596496 3.708576 3.708576 3.556383
3.596496 3.628076
 12   13   14   15   16   17   18   19   20   21   22
3.675810 3.668462 3.700801 3.594724 3.628076 3.628076 3.732346 3.596496 3.596496
3.732346 3.571540
 23   24   25   26   27   28   29   30   31   32   33
3.611654 3.717224 3.715923 3.724998 3.505576 3.571540 3.602499 3.732346 3.556383
3.596496 3.732346
 34   35   36   37   38   39   40   41   42   43   44
3.715923 3.505576 3.571540 3.611654 3.660688 3.708576 3.612919 3.490419 3.717224
3.660688 3.675810
 45   46   47   48   49   50   51   52   53   54   55
3.628076 3.708576 3.732346 3.612919 3.660688 3.612919 3.628076 3.700801 3.594724
3.708576 3.717224
 56   57   58   59   60   61   62   63   64   65   66
3.628076 3.628076 3.628076 3.609846 3.700801 3.668462 3.611654 3.724998 3.505576
3.717224 3.715923
 67   68   69   70   71   72   73   74   75   76   77
3.505576 3.596496 3.732346 3.717224 3.700801 3.596496 3.609846 3.724998 3.594724
3.611654 3.571540
 78   79   80   81   82   83   84   85   86   87   88
3.708576 3.715923 3.715923 3.717224 3.609846 3.609846 3.724998 3.708576 3.490419
3.717224 3.490419
 89   90   91   92   93   94   95   96   97   98   99
3.732346 3.602499 3.668462 3.708576 3.717224 3.596496 3.715923 3.612919 3.628076
3.628076 3.732346
 100
3.609846

```

```
> plot(proje5$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje5$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Doğrusallığı istediğimiz gibi olmayan bir grafik daha elde ettik.

```
> AIC(proje5)
[1] 655.1403
```

++AIC değerimiz her testte düşmeye devam ediyor. Son kombinasyona kadar değerine bakacağız.

```
> anova(proje5, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			99	108.214	
vehicle_type2	3	6.8668	96	101.347	0.07627 .

```
cities2    4 10.0521    92  91.295 0.03956 *
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++Diğer çıktıda olduğu gibi bunda da “cities2” değişkeni 0,05 ten küçük. Ho red. Model anlamsız.

BÖLÜM-7

```
> proje6 <- glm(hasar_sayisi ~ gender2 + cities2 ,fam = poisson(link = log))
> summary(proje6)
```

Call:

```
glm(formula = hasar_sayisi ~ gender2 + cities2, family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.09603	-0.65020	0.07326	0.64411	2.06287

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.66451	0.03520	104.113	<2e-16 ***
gender22	0.06475	0.03274	1.978	0.048 *
cities22	-0.10089	0.04785	-2.108	0.035 *
cities23	-0.01372	0.04986	-0.275	0.783
cities24	-0.01270	0.04940	-0.257	0.797
cities25	-0.11853	0.05095	-2.326	0.020 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.214 on 99 degrees of freedom
Residual deviance: 94.427 on 94 degrees of freedom
AIC: 654.27

Number of Fisher Scoring iterations: 4

minimum değerim; -2.09603

1.çeyreklik; -0.65020

ortanca değer; 0.07326

3. çeyreklik; 0.64411

maksimum değerim; 2.06287

++Başlangıç değeri; 3.66451 tir.

++”AIC” değerini en son yorumlayacağım.

++ “names” komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

> names(proje6)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

> proje6\$coef

```
(Intercept) gender22 cities22 cities23 cities24 cities25
3.66451113 0.06474543 -0.10089219 -0.01371748 -0.01270366 -0.11852919
```

++ $Y = 3.66451113 + 0.06474543 (x_1) + -0.10089219(x_2) + -0.01371748 (x_3) + -0.01270366(x_4) + -0.11852919 (x_5)$

modelini kurabiliriz.

> proje6\$fitted.values

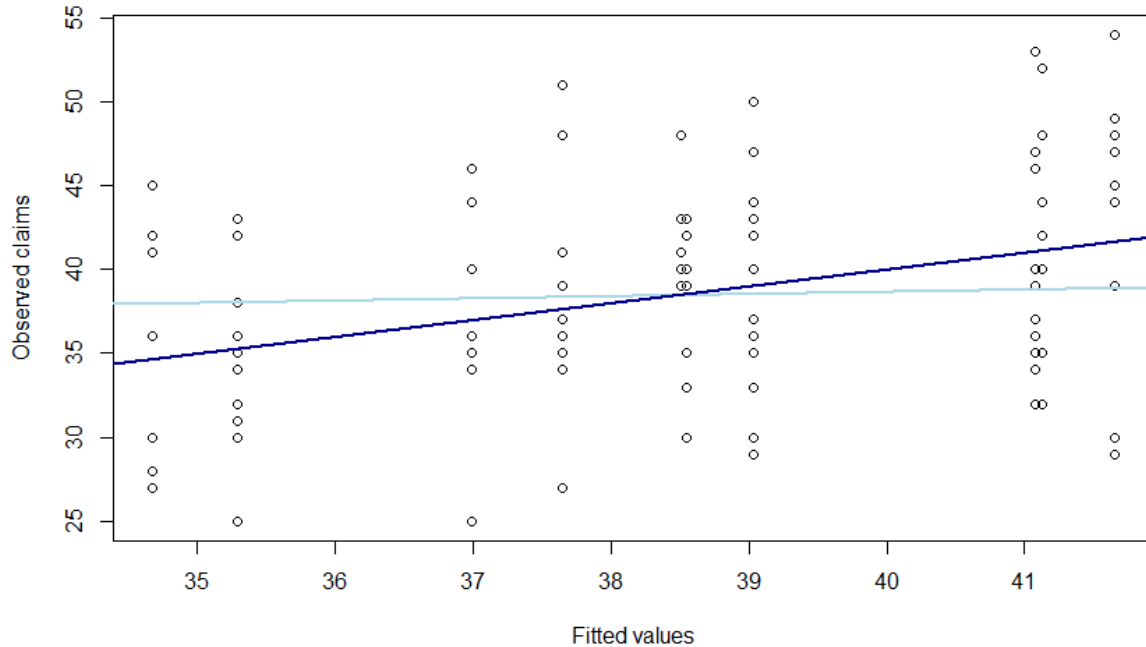
```
 1    2    3    4    5    6    7    8    9   10   11
39.03705 41.64813 34.67372 41.08073 41.64813 36.99295 38.50521 41.08073 36.99295
36.99295 35.29068
12   13   14   15   16   17   18   19   20   21   22
39.03705 41.08073 41.12240 38.54427 35.29068 37.65118 41.64813 34.67372 34.67372
39.03705 35.29068
23   24   25   26   27   28   29   30   31   32   33
37.65118 38.54427 39.03705 38.50521 37.65118 35.29068 38.50521 39.03705 36.99295
36.99295 39.03705
34   35   36   37   38   39   40   41   42   43   44
39.03705 35.29068 35.29068 37.65118 38.54427 41.08073 36.99295 34.67372 38.54427
38.54427 41.64813
45   46   47   48   49   50   51   52   53   54   55
37.65118 41.08073 39.03705 36.99295 41.12240 36.99295 37.65118 38.54427 41.12240
41.08073 38.54427
56   57   58   59   60   61   62   63   64   65   66
35.29068 37.65118 37.65118 41.64813 38.54427 41.08073 35.29068 41.08073 35.29068
41.12240 39.03705
67   68   69   70   71   72   73   74   75   76   77
35.29068 36.99295 41.64813 41.12240 41.12240 34.67372 39.03705 38.50521 41.12240
37.65118 35.29068
78   79   80   81   82   83   84   85   86   87   88
```

```
38.50521 39.03705 39.03705 41.12240 39.03705 41.64813 41.08073 41.08073 34.67372
38.54427 34.67372
  89   90   91   92   93   94   95   96   97   98   99
41.64813 41.08073 41.08073 38.50521 41.12240 34.67372 39.03705 34.67372 37.65118
35.29068 41.64813
  100
41.64813
```

> proje6\$linear.predictors

```
  1    2    3    4    5    6    7    8    9   10   11
3.664511 3.729257 3.545982 3.715539 3.729257 3.610727 3.650794 3.715539 3.610727
3.610727 3.563619
 12   13   14   15   16   17   18   19   20   21   22
3.664511 3.715539 3.716553 3.651807 3.563619 3.628364 3.729257 3.545982 3.545982
3.664511 3.563619
 23   24   25   26   27   28   29   30   31   32   33
3.628364 3.651807 3.664511 3.650794 3.628364 3.563619 3.650794 3.664511 3.610727
3.610727 3.664511
 34   35   36   37   38   39   40   41   42   43   44
3.664511 3.563619 3.563619 3.628364 3.651807 3.715539 3.610727 3.545982 3.651807
3.651807 3.729257
 45   46   47   48   49   50   51   52   53   54   55
3.628364 3.715539 3.664511 3.610727 3.716553 3.610727 3.628364 3.651807 3.716553
3.715539 3.651807
 56   57   58   59   60   61   62   63   64   65   66
3.563619 3.628364 3.628364 3.729257 3.651807 3.715539 3.563619 3.715539 3.563619
3.716553 3.664511
 67   68   69   70   71   72   73   74   75   76   77
3.563619 3.610727 3.729257 3.716553 3.716553 3.545982 3.664511 3.650794 3.716553
3.628364 3.563619
 78   79   80   81   82   83   84   85   86   87   88
3.650794 3.664511 3.664511 3.716553 3.664511 3.729257 3.715539 3.715539 3.545982
3.651807 3.545982
 89   90   91   92   93   94   95   96   97   98   99
3.729257 3.715539 3.715539 3.650794 3.716553 3.545982 3.664511 3.545982 3.628364
3.563619 3.729257
 100
3.729257
```

```
> plot(proje6$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje6$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Saçılımın fazla olduğu bir grafik elde ettik.

```
> AIC(proje6)
[1] 654.2727
```

++AIC değerimiz, diğer testlerdeki değer ile benzerlik göstermekte.

```
> anova(proje6, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			99	108.214	
gender2 1	4.3243		98	103.889	0.03757 *
cities2 4	9.4622		94	94.427	0.05053 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++"gender2" değişkeninde, $\alpha < 0,05$. Bu sebeple H_0 red, model anlamlı değildir.

BÖLÜM-8

```
> proje7 <- glm(hasar_sayisi ~ non_damaged2 + cities2 + gender2 ,fam = poisson(link = log))
> summary(proje7)
```

Call:

```
glm(formula = hasar_sayisi ~ non_damaged2 + cities2 + gender2,
     family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.26901	-0.66258	0.03288	0.64572	1.91177

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.6567213	0.0437855	83.514	<2e-16 ***
non_damaged22	0.0445438	0.0403700	1.103	0.2699
non_damaged23	0.0008075	0.0404816	0.020	0.9841
cities22	-0.1093889	0.0488614	-2.239	0.0252 *
cities23	-0.0278866	0.0513456	-0.543	0.5870
cities24	-0.0152078	0.0499915	-0.304	0.7610
cities25	-0.1185460	0.0511053	-2.320	0.0204 *
gender22	0.0597517	0.0331294	1.804	0.0713 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.214 on 99 degrees of freedom
Residual deviance: 92.901 on 92 degrees of freedom
AIC: 656.75

Number of Fisher Scoring iterations: 4

minimum değerim; -2.26901
1.çeyreklik; -0.66258
ortanca değer; 0.03288
3. çeyreklik; 0.64572
maksimum değerim; 1.91177

++Başlangıç değeri; 3.6567213 tir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.


```
> names(proje7)
```

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

```
> proje7$coef
```

```
(Intercept) non_damaged22 non_damaged23 cities22 cities23 cities24 cities25
3.6567213104 0.0445438053 0.0008074828 -0.1093888899 -0.0278866279
-0.0152077884 -0.1185459729
gender22
0.0597517451
```

$++Y = 3.6567213104 + 0.0445438053 (x1) + 0.0008074828(x2) + -0.1093888899 (x3) + -0.0278866279 (x4) + -0.0152077884 (x5) + -0.1185459729(x6) + 0.0597517451(x7)$

modelini kurabiliriz.

```
> proje7$fitted.values
```

```
 1    2    3    4    5    6    7    8    9   10   11
40.49851 41.15233 34.40409 41.80978 41.11911 36.55195 39.38474 41.80978 36.52245
36.52245 34.72057
12   13   14   15   16   17   18   19   20   21   22
38.76543 39.98828 40.49851 38.14953 34.72057 36.88820 41.11911 35.97122 34.43188
38.76543 34.72057
23   24   25   26   27   28   29   30   31   32   33
38.53735 38.14953 38.76543 37.69932 36.88820 34.72057 39.38474 38.76543 36.55195
38.18607 38.76543
34   35   36   37   38   39   40   41   42   43   44
40.49851 36.30212 34.72057 36.85842 38.18035 41.80978 38.18607 34.43188 38.14953
38.18035 41.15233
45   46   47   48   49   50   51   52   53   54   55
36.88820 41.80978 38.76543 36.55195 40.49851 36.52245 36.85842 38.14953 42.34325
40.02058 38.18035
56   57   58   59   60   61   62   63   64   65   66
```

```

36.30212 38.53735 36.85842 42.99212 39.88727 41.80978 36.30212 41.80978 34.74862
40.53123 38.73414
  67   68   69   70   71   72   73   74   75   76   77
34.74862 36.52245 41.15233 40.49851 42.34325 35.97122 38.76543 37.69932 40.49851
38.53735 36.30212
  78   79   80   81   82   83   84   85   86   87   88
37.69932 38.76543 38.73414 40.53123 38.73414 42.99212 39.98828 41.80978 34.40409
39.88727 34.40409
  89   90   91   92   93   94   95   96   97   98   99
41.11911 39.98828 41.80978 37.66889 42.34325 34.43188 38.73414 34.43188 38.53735
36.30212 42.99212
  100
41.15233

```

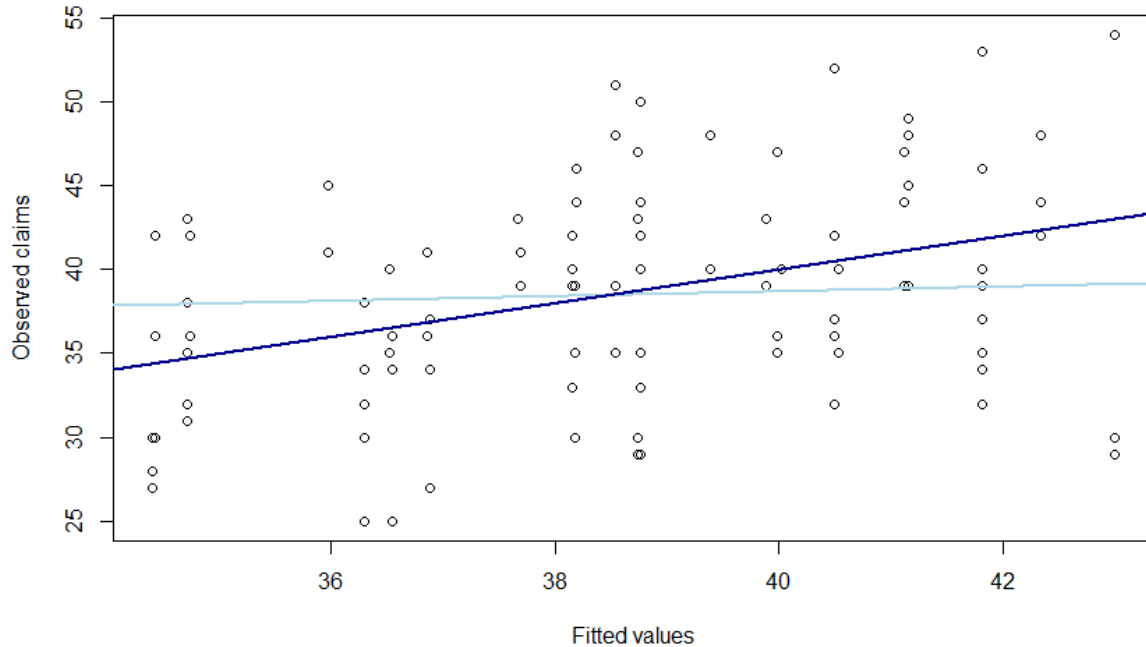
> proje7\$linear.predictors

```

  1   2   3   4   5   6   7   8   9  10  11
3.701265 3.717281 3.538175 3.733130 3.716473 3.598735 3.673378 3.733130 3.597927
3.597927 3.547332
 12  13  14  15  16  17  18  19  20  21  22
3.657529 3.688586 3.701265 3.641514 3.547332 3.607892 3.716473 3.582719 3.538983
3.657529 3.547332
 23  24  25  26  27  28  29  30  31  32  33
3.651628 3.641514 3.657529 3.629642 3.607892 3.547332 3.673378 3.657529 3.598735
3.642471 3.657529
 34  35  36  37  38  39  40  41  42  43  44
3.701265 3.591876 3.547332 3.607084 3.642321 3.733130 3.642471 3.538983 3.641514
3.642321 3.717281
 45  46  47  48  49  50  51  52  53  54  55
3.607892 3.733130 3.657529 3.598735 3.701265 3.597927 3.607084 3.641514 3.745809
3.689394 3.642321
 56  57  58  59  60  61  62  63  64  65  66
3.591876 3.651628 3.607084 3.761017 3.686057 3.733130 3.591876 3.733130 3.548140
3.702073 3.656721
 67  68  69  70  71  72  73  74  75  76  77
3.548140 3.597927 3.717281 3.701265 3.745809 3.582719 3.657529 3.629642 3.701265
3.651628 3.591876
 78  79  80  81  82  83  84  85  86  87  88
3.629642 3.657529 3.656721 3.702073 3.656721 3.761017 3.688586 3.733130 3.538175
3.686057 3.538175
 89  90  91  92  93  94  95  96  97  98  99
3.716473 3.688586 3.733130 3.628835 3.745809 3.538983 3.656721 3.538983 3.651628
3.591876 3.761017
 100
3.717281

```

```
> plot(proje7$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje7$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Ürettiğimiz grafik hala çok dağınık seyrediyor.

```
> AIC(proje7)
[1] 656.7464
```

++AIC değerimiz beklentiden yüksek çıktı.

```
> anova(proje7, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			99	108.214	
non_damaged2	2	2.2538	97	105.960	0.32404
cities2	4	9.8055	93	96.154	0.04383 *
gender2	1	3.2535	92	92.901	0.07127 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++ "cities2" değişkenimiz $\alpha < 0,05$ sınırını aldı. Ancak çok küçük bir fark olduğundan dolayı bu modeli reddetmek doğru olmayacak. Bu yüzden H_0 reddedilemez, model anlamlıdır yorumunu yapmak istiyorum.

BÖLÜM-9

```
> proje8 <- glm(hasar_sayisi ~ non_damaged2 + gender2 + vehicle_type2, fam =  
poisson(link = log))  
> summary(proje8)
```

Call:

```
glm(formula = hasar_sayisi ~ non_damaged2 + gender2 + vehicle_type2,  
family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.50237	-0.71157	-0.01683	0.71879	2.08785

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.638222	0.040858	89.045	<2e-16 ***
non_damaged22	0.043781	0.039906	1.097	0.2726
non_damaged23	0.007601	0.040107	0.190	0.8497
gender22	0.055357	0.033050	1.675	0.0939 .
vehicle_type22	-0.020318	0.040293	-0.504	0.6141
vehicle_type23	-0.046540	0.047977	-0.970	0.3320
vehicle_type24	-0.115195	0.049384	-2.333	0.0197 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.214 on 99 degrees of freedom
Residual deviance: 96.488 on 93 degrees of freedom
AIC: 658.33

Number of Fisher Scoring iterations: 4

++minimum değerim; -2.50237

1.çeyreklik; -0.71157

ortanca değer; -0.01683

3. çeyreklik; 0.71879

maksimum değerim; 2.08785

++Başlangıç değeri; 3.638222 dir.

++"AIC" değerini en son yorumlayacağım.

++ "names" komutu ile GLM hesabında araştırabileceğimiz konu başlıklarını görmüş olduk.

++Serbestlik derecesi (n-1) olduğundan 99 sayısını elde ettik.

> names(proje8)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

> proje8\$coef

```
(Intercept) non_damaged22 non_damaged23 gender22 vehicle_type22
vehicle_type23
 3.638222044 0.043780944 0.007600589 0.055357247 -0.020318262
-0.046539845
vehicle_type24
-0.115195496
```

++ $Y = 3.638222044 + 0.043780944 (x_1) + 0.007600589 (x_2) + 0.055357247 (x_3) +$
 $-0.020318262 (x_4) - 0.046539845 (x_5) + -0.115195496(x_6)$

modelini kurabiliriz.

> proje8\$fitted.values

```
 1    2    3    4    5    6    7    8    9   10   11
37.91941 38.65360 37.25938 41.14251 38.36093 39.68057 38.92687 41.14251 38.36093
39.38012 38.02417
12   13   14   15   16   17   18   19   20   21   22
36.57200 38.36093 39.38012 33.88683 38.02417 40.49506 40.18844 38.92687 37.54365
38.31428 36.29508
23   24   25   26   27   28   29   30   31   32   33
41.14251 38.02417 37.54365 38.31428 36.08887 36.29508 35.40339 38.31428 38.65360
41.14251 38.31428
34   35   36   37   38   39   40   41   42   43   44
38.92687 35.40339 36.29508 39.38012 36.57200 41.14251 41.98701 34.14537 38.02417
36.57200 38.65360
45   46   47   48   49   50   51   52   53   54   55
```

```

40.49506 41.14251 38.31428 40.49506 38.36093 40.18844 40.18844 37.25938 37.41848
39.68057 38.31428
  56    57    58    59    60    61    62    63    64    65    66
39.72588 41.98701 40.18844 37.41848 38.92687 40.07771 38.92687 41.98701 34.14537
40.49506 37.25938
  67    68    69    70    71    72    73    74    75    76    77
34.14537 39.38012 40.49506 40.18844 41.14251 38.92687 34.14537 38.31428 35.81561
41.14251 37.91941
  78    79    80    81    82    83    84    85    86    87    88
37.54365 37.54365 37.25938 40.49506 33.88683 37.41848 40.18844 41.14251 33.88683
39.72588 33.88683
  89    90    91    92    93    94    95    96    97    98    99
40.18844 35.81561 40.07771 37.25938 41.98701 37.54365 37.25938 38.31428 41.98701
39.72588 41.98701
 100
36.08887

```

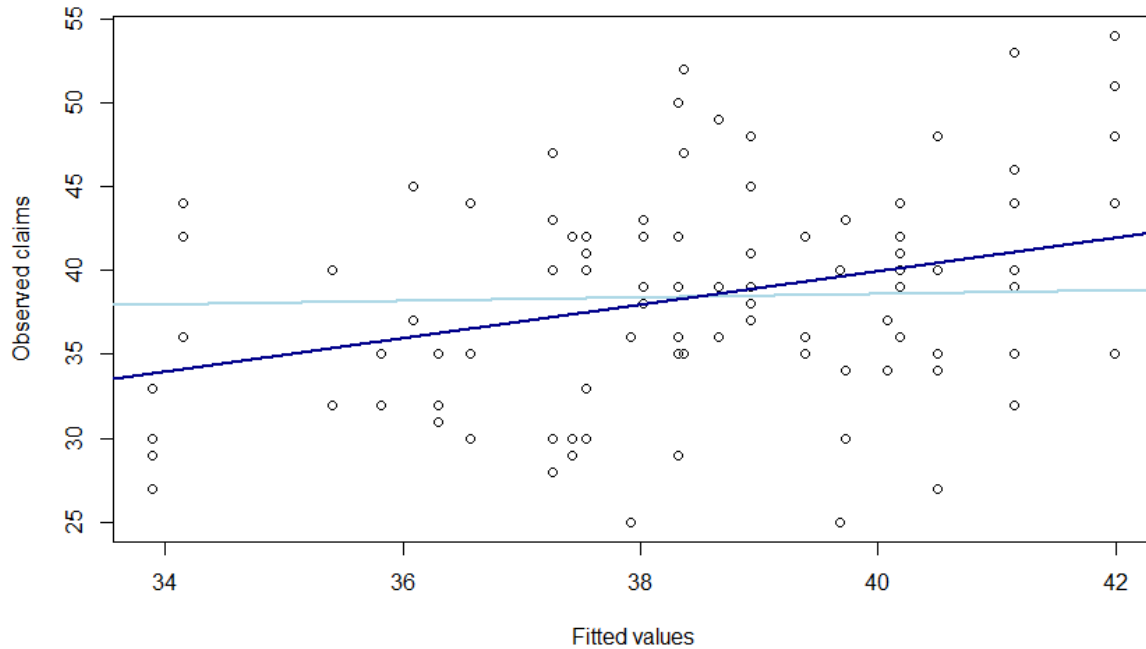
> proje8\$linear.predictors

```

  1    2    3    4    5    6    7    8    9   10   11
3.635463 3.654640 3.617904 3.717042 3.647039 3.680862 3.661685 3.717042 3.647039
3.673261 3.638222
 12   13   14   15   16   17   18   19   20   21   22
3.599283 3.647039 3.673261 3.523027 3.638222 3.701180 3.693579 3.661685 3.625504
3.645823 3.591682
 23   24   25   26   27   28   29   30   31   32   33
3.717042 3.638222 3.625504 3.645823 3.585984 3.591682 3.566807 3.645823 3.654640
3.717042 3.645823
 34   35   36   37   38   39   40   41   42   43   44
3.661685 3.566807 3.591682 3.673261 3.599283 3.717042 3.737360 3.530627 3.638222
3.599283 3.654640
 45   46   47   48   49   50   51   52   53   54   55
3.701180 3.717042 3.645823 3.701180 3.647039 3.693579 3.693579 3.617904 3.622165
3.680862 3.645823
 56   57   58   59   60   61   62   63   64   65   66
3.682003 3.737360 3.693579 3.622165 3.661685 3.690820 3.661685 3.737360 3.530627
3.701180 3.617904
 67   68   69   70   71   72   73   74   75   76   77
3.530627 3.673261 3.701180 3.693579 3.717042 3.661685 3.530627 3.645823 3.578384
3.717042 3.635463
 78   79   80   81   82   83   84   85   86   87   88
3.625504 3.625504 3.617904 3.701180 3.523027 3.622165 3.693579 3.717042 3.523027
3.682003 3.523027
 89   90   91   92   93   94   95   96   97   98   99
3.693579 3.578384 3.690820 3.617904 3.737360 3.625504 3.617904 3.645823 3.737360
3.682003 3.737360
 100
3.585984

```

```
> plot(proje8$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje8$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Lacivert doğru çevresinde biraz daha fazla daire gözlemliyoruz. Bu iyi bir şey ancak saçılım hala çok fazla.

```
> AIC(proje8)
[1] 658.3336
```

++AIC değerimiz çok yüksek çıktı.

```
> anova(proje8, test="Chisq")
Analysis of Deviance Table
```

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			99	108.214	
non_damaged2	2	2.2538	97	105.960	0.3240
gender2	1	3.5823	96	102.378	0.0584 .
vehicle_type2	3	5.8895	93	96.488	0.1171

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++ Alfa > 0.05, Ho reddedilemez. Model anlamlıdır.

BÖLÜM-10

```
> proje9 <- glm(hasar_sayisi ~ cities2 + gender2 + vehicle_type2 ,fam = poisson(link = log))
> summary(proje9)
```

Call:

```
glm(formula = hasar_sayisi ~ cities2 + gender2 + vehicle_type2,
     family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.20042	-0.57578	-0.08137	0.69689	1.87626

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.70272	0.04263	86.857	<2e-16 ***
cities22	-0.10656	0.04805	-2.218	0.0266 *
cities23	-0.02418	0.05091	-0.475	0.6348
cities24	-0.01980	0.04955	-0.400	0.6895
cities25	-0.12694	0.05190	-2.446	0.0144 *
gender22	0.06092	0.03301	1.845	0.0650 .
vehicle_type22	-0.00778	0.04125	-0.189	0.8504
vehicle_type23	-0.05469	0.04810	-1.137	0.2556
vehicle_type24	-0.11449	0.04944	-2.316	0.0206 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.214 on 99 degrees of freedom
Residual deviance: 87.888 on 91 degrees of freedom
AIC: 653.73

Number of Fisher Scoring iterations: 4

++minimum değerim; -2.20042
1.çeyreklik; -0.57578
ortanca değer; -0.08137
3. çeyreklik; 0.69689
maksimum değerim; 1.87626

++Başlangıç değeri; 3.70272 dir.


```
> names(proje9)
```

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

```
> proje9$coef
```

```
(Intercept) cities22 cities23 cities24 cities25 gender22
3.702715631 -0.106563455 -0.024181213 -0.019798247 -0.126939622
0.060917912
vehicle_type22 vehicle_type23 vehicle_type24
-0.007780432 -0.054689412 -0.114493019
```

```
++Y = 3.702715631 - 0.106563455(x1) + -0.024181213 (x2) + -0.019798247 (x3) +
-0.126939622 (x4) + 0.060917912 (x5) + -0.007780432 (x6) + -0.054689412(x7) +
-0.114493019(x8)
```

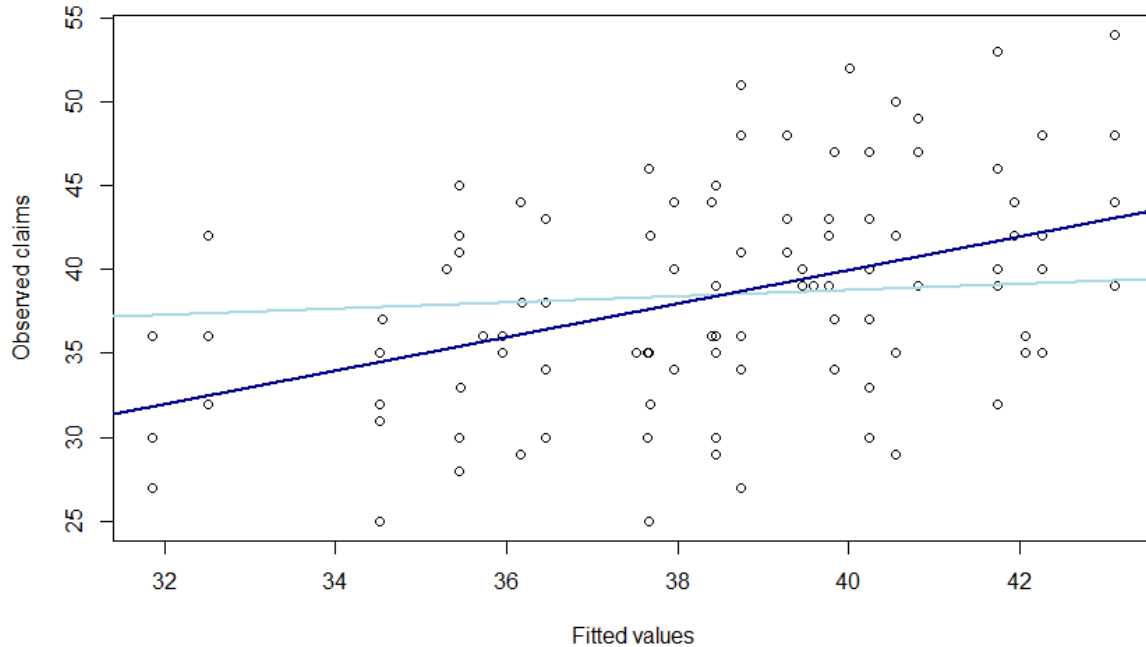
modelini kurabiliriz.

```
> proje9$fitted.values
```

```
1 2 3 4 5 6 7 8 9 10 11
38.39880 40.81069 35.44547 41.74885 40.81069 37.67186 39.28151 41.74885 35.94552
37.67186 36.45768
12 13 14 15 16 17 18 19 20 21 22
38.39880 39.83568 41.93224 35.46068 36.45768 38.74765 43.10476 35.44547 35.44547
40.55729 34.51737
23 24 25 26 27 28 29 30 31 32 33
38.44735 39.76223 40.24296 39.58833 34.55586 34.51737 35.30559 40.55729 35.94552
37.67186 40.55729
34 35 36 37 38 39 40 41 42 43 44
40.24296 32.51362 34.51737 38.44735 37.64605 41.74885 37.96611 31.85782 39.76223
37.64605 40.81069
45 46 47 48 49 50 51 52 53 54 55
38.74765 41.74885 40.55729 37.96611 40.01066 37.96611 38.74765 39.45406 37.68802
41.74885 39.76223
56 57 58 59 60 61 62 63 64 65 66
36.45768 38.74765 38.74765 38.44161 39.45406 39.83568 36.17513 42.07494 32.51362
42.25976 40.24296
```



```
> plot(proje9$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje9$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)
```



++Eğim alanına giren çok fazla değer yok. Yine saçılımın fazla olmasıyla birlikte, tek bir yerde yığılma gerçekleşmemiş.

```
> AIC(proje9)
```

```
[1] 653.7337
```

++Grafik ile birlikte AIC değerine baktığımızda, ilk test dışındakilerden daha iyi senaryo gözlemliyoruz.

```
> anova(proje9, test="Chisq")
```

Analysis of Deviance Table

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			99	108.214	
cities2	4	9.8733	95	98.340	0.04262 *
gender2	1	3.9132	94	94.427	0.04791 *

```
vehicle_type2 3 6.5390 91 87.888 0.08814 .
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++Her iki değişkende de $\alpha < 0,05$. Ho Red, model anlamsızdır.

BÖLÜM-11

```
> proje10 <- glm(hasar_sayisi ~ non_damaged2 + cities2 + vehicle_type2 ,fam =  
poisson(link = log))  
> summary(proje10)
```

Call:

```
glm(formula = hasar_sayisi ~ non_damaged2 + cities2 + vehicle_type2,  
family = poisson(link = log))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.06094	-0.57591	-0.00941	0.64638	2.04301

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.73165	0.04792	77.875	<2e-16 ***
non_damaged22	0.04739	0.04067	1.165	0.2439
non_damaged23	-0.01078	0.04067	-0.265	0.7910
cities22	-0.11825	0.04939	-2.394	0.0167 *
cities23	-0.02428	0.05140	-0.472	0.6367
cities24	-0.02155	0.05021	-0.429	0.6678
cities25	-0.11877	0.05184	-2.291	0.0220 *
vehicle_type22	-0.02863	0.04198	-0.682	0.4953
vehicle_type23	-0.05621	0.04821	-1.166	0.2437
vehicle_type24	-0.12642	0.04940	-2.559	0.0105 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 108.21 on 99 degrees of freedom
Residual deviance: 89.14 on 90 degrees of freedom
AIC: 656.99

Number of Fisher Scoring iterations: 4

++minimum değerim; -2.06094
1.çeyreklik; -0.57591

ortanca değeri; -0.00941
3. çeyreklik; 0.64638
maksimum değeri; 2.04301

++Başlangıç değeri; 3.73165 tir.

> names(proje10)

```
[1] "coefficients" "residuals" "fitted.values" "effects"
[5] "R" "rank" "qr" "family"
[9] "linear.predictors" "deviance" "aic" "null.deviance"
[13] "iter" "weights" "prior.weights" "df.residual"
[17] "df.null" "y" "converged" "boundary"
[21] "model" "call" "formula" "terms"
[25] "data" "offset" "control" "method"
[29] "contrasts" "xlevels"
```

> proje10\$coef

```
(Intercept) non_damaged22 non_damaged23 cities22 cities23 cities24
3.73165047 0.04739248 -0.01077822 -0.11824774 -0.02427530 -0.02154968
cities25 vehicle_type22 vehicle_type23 vehicle_type24
-0.11876581 -0.02862675 -0.05620567 -0.12642135
```

++Y = 3.73165047 + 0.04739248 (x1) + -0.01077822 (x2) + -0.11824774 (x3) +
-0.02427530 (x4) + -0.02154968(x5) + -0.11876581(x6) + -0.02862675(x7) +
-0.05620567(x8) + -0.12642135(x9)

modelini kurduk.

> proje10\$fitted.values

```
1 2 3 4 5 6 7 8 9 10 11
41.38164 39.04312 36.02661 41.51857 39.46621 35.64040 41.51857 41.51857 35.04661
36.02661 37.09205
12 13 14 15 16 17 18 19 20 21 22
39.04312 38.51969 39.70488 36.00578 37.09205 36.69441 41.74796 37.77511 35.64040
41.30040 35.06477
23 24 25 26 27 28 29 30 31 32 33
37.79468 40.85792 40.13487 40.30989 32.33671 35.06477 37.65050 41.30040 34.67090
37.77511 41.30040
34 35 36 37 38 39 40 41 42 43 44
```

```

42.53878 34.27355 35.06477 36.04528 38.21075 41.51857 38.87211 32.31996 40.85792
38.21075 39.04312
  45   46   47   48   49   50   51   52   53   54   55
36.69441 41.51857 41.30040 36.67541 38.62482 37.07284 37.09205 39.70488 37.75326
39.17231 40.41991
  56   57   58   59   60   61   62   63   64   65   66
38.89226 38.89226 37.09205 38.57566 41.63189 40.38918 37.79468 42.72429 32.33671
40.41991 40.56979
  67   68   69   70   71   72   73   74   75   76   77
32.33671 36.02661 41.30040 40.85792 41.63189 37.77511 36.39571 40.30989 36.00578
37.79468 36.76659
  78   79   80   81   82   83   84   85   86   87   88
39.17231 40.13487 40.56979 40.41991 36.79011 38.57566 40.74671 41.51857 32.67020
42.84090 32.67020
  89   90   91   92   93   94   95   96   97   98   99
41.74796 35.90777 40.38918 39.59680 42.84090 35.64040 40.56979 36.67541 38.89226
38.89226 43.77413
  100
36.39571

```

> proje10\$linear.predictors

```

  1    2    3    4    5    6    7    8    9   10   11
3.722837 3.664667 3.584258 3.726141 3.675445 3.573480 3.726141 3.726141 3.556679
3.584258 3.613403
 12   13   14   15   16   17   18   19   20   21   22
3.664667 3.651169 3.681474 3.583679 3.613403 3.602625 3.731650 3.631650 3.573480
3.720872 3.557197
 23   24   25   26   27   28   29   30   31   32   33
3.632168 3.710101 3.692246 3.696597 3.476203 3.557197 3.628346 3.720872 3.545901
3.631650 3.720872
 34   35   36   37   38   39   40   41   42   43   44
3.750416 3.534374 3.557197 3.584776 3.643117 3.726141 3.660277 3.475685 3.710101
3.643117 3.664667
 45   46   47   48   49   50   51   52   53   54   55
3.602625 3.726141 3.720872 3.602106 3.653895 3.612885 3.613403 3.681474 3.631072
3.667970 3.699323
 56   57   58   59   60   61   62   63   64   65   66
3.660795 3.660795 3.613403 3.652622 3.728867 3.698562 3.632168 3.754768 3.476203
3.699323 3.703024
 67   68   69   70   71   72   73   74   75   76   77
3.476203 3.584258 3.720872 3.710101 3.728867 3.631650 3.594451 3.696597 3.583679
3.632168 3.604590
 78   79   80   81   82   83   84   85   86   87   88

```

```

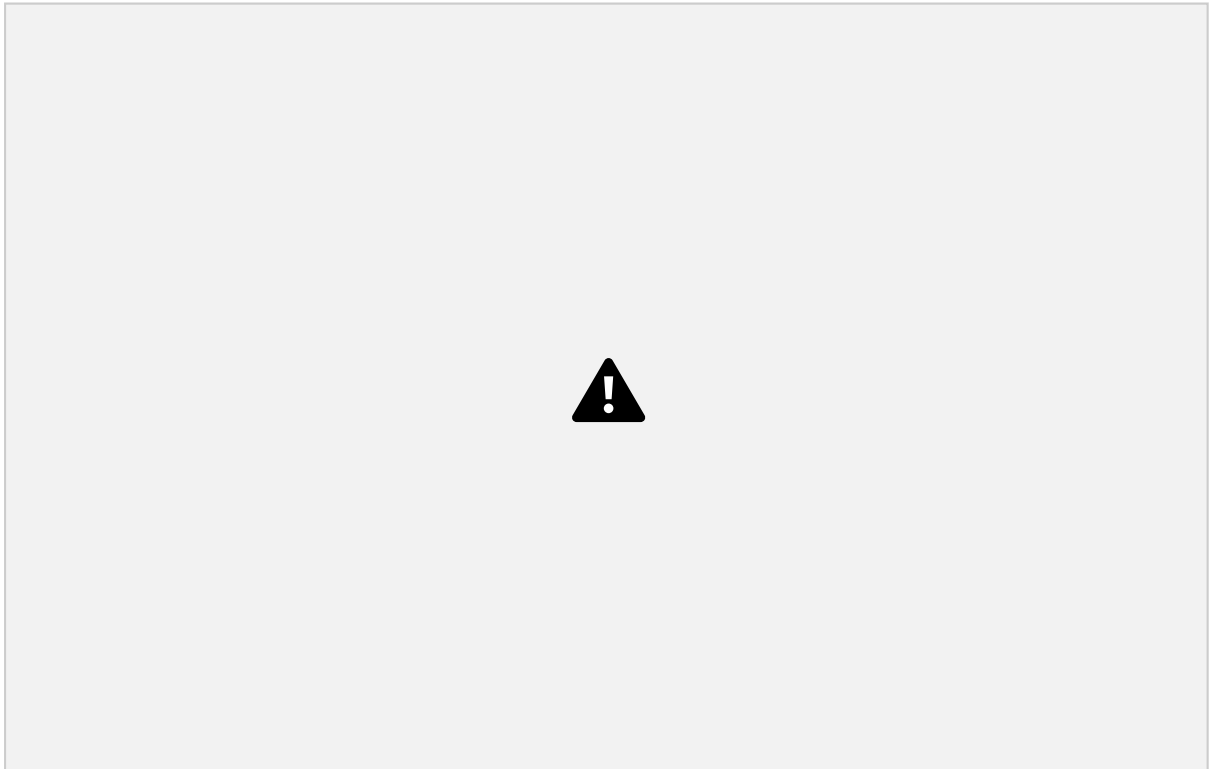
3.667970 3.692246 3.703024 3.699323 3.605229 3.652622 3.707375 3.726141 3.486463
3.757493 3.486463
   89   90   91   92   93   94   95   96   97   98   99
3.731650 3.580954 3.698562 3.678748 3.757493 3.573480 3.703024 3.602106 3.660795
3.660795 3.779043
   100
3.594451

```

```

> plot(proje10$fitted.values, hasar_sayisi, xlab = "Fitted values", ylab = "Observed claims")
> abline(lm(proje10$fitted ~ hasar_sayisi), col="light blue", lwd=2)
> abline(0, 1, col = "dark blue", lwd=2)

```



++Görüşüme göre ilki dışında elde ettiğimiz en iyi grafik budur. Ancak saçılım hala çok fazladır.

```
> AIC(proje10)
```

```
[1] 656.9851
```

++AIC değerim yine yüksek çıktı.

```
> anova(proje10, test="Chisq")
```

Analysis of Deviance Table

Model: poisson, link: log

Response: hasar_sayisi

Terms added sequentially (first to last)

		Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL		99	108.214			
non_damaged2	2	2.2538	97	105.960	0.32404	
cities2	4	9.8055	93	96.154	0.04383	*
vehicle_type2	3	7.0148	90	89.140	0.07143	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

++"cities2" değişkeni, $\alpha < 0.05$ çıktı. H_0 red, model anlamsızdır.

YORUM BÖLÜMÜ

Hasar sayısı, rasgele üretildiğinden dolayı grafiksel olarak pek uygun olmadı. Ancak bu varsayım altında konuşmak gerekirse ;

*Değişkenlerimi bir arada düşünmemin doğru olacağı görüşündeyim. Değişkenlerim birbirini etkiliyor ve ikili kombinasyonlarda bağılantısız sonuçlar veriyor.

*Ayrı ayrı değerlendirmeler sağlıklı sonuca ulaşmamda engel oluşturuyor. Bir risk sınıflandırması yapmam durumunda bu değişkenler hesaplamalarımda yanlış sonuca ulaşmama sebep olabilir.

*Eğer bir sigorta şirketi aktüeri olsaydım, verilen bilgiler ve analizim dahilinde bu sigorta türünden kaçınırdım. Çünkü tutarsız sonuçlar beni zarara sokabilir. Yüksek prim uygulamasına giderdim. Hem kendimi riske atmazdım, hem de doğal olarak müşteriler bu sigorta türünden kaçınmış olurdu.

*Uzun kuyruklu ve katastrofik olmayan riskler için kullanmak daha doğru olurdu. Eğer ki kısa kuyruklu sigorta türünde kullanırsam, göz göre göre şirketi zarara sokardım. Onun dışında uzun kuyruklu ve katastrofik riskler için kullansaydım, ödeyeceğim tazminatlar çok fazla olurdu.

*Eğer ki şansım olsaydı modele uymayan değişkenleri çıkartırdım. Örneğin risk sınıflandırması yaptığımda, araç türü ile cinsiyet birbirine karşı tutarsız ise, o araç türüne sahip cinsiyetteki müşterilere poliçe satmazdım.

Ya da x şehrinde cinsiyet bakımından hasar fazla gelecekse, radikal karar alıp o şehirde sigorta kolu açmazdım. Herhangi bir cinsiyete poliçe satılmaması daha büyük tepkiye sebep olabilirdi. Ya da primi çok yüksek tutardım.

*Elbette tüm bunları gözlemlerken Büyük Sayılar Kanunu'nu da göz önünde bulundurmamız gerekirdi. Ancak elimizdeki veriler ışığında izlenimlerim bunlardır.

*Sonuç olarak eğer ki elimdeki verilere göre referans alacaksam, Bölüm-1 de bulunan tüm değişkenlerin olduğu teste göre hareket etmem en doğru tercih olacaktır.