ABD'de İkinci El Araba Alma Kılavuzu

Ş. İlker Birbil

3 Mart 2016

Böyle bir başlık okuyucu kaçırır. Farkındayım. Aman kaçmayın. Ulvi bir mesele için buradayız.

Bu ikinci el araba işi ABD'de okuyacaklar için büyük bir hadisedir. Üniversiteye yazılır, bir eve yerleşirsiniz. Sonra bir bakmışsınız etrafta bakkal çakkal yok. Ekmeğiniz bitse, en yakın markete gitmek için araba gerekir. Neyse ki ABD'de arabalar, Türkiye'deki gibi öyle yanlarına yaklaşılmayacak fiyatlarda değiller. Hele ikinci ellere bakarsanız -ki öğrenciliğin raconudur- cebinizden çıkacak para oldukça makul olabilir.

O makul meblağı denkleştirince araba avı başlar. Dört tekerden beklenen aşağı yukarı bellidir: "Yolda bırakmasın, ucuz olsun; benim olsun." Öte yandan araba seçmek de öyle kolay bir iş değildir. Vitesli mi olsun, otomatik mi? Yaşına mı bakmalı, yoksa kaç kilometrede olduğuna mı? Sorular, sorular...

Tamam bir kılavuz yazayım. Fakat bir [Bol Bilim](http://www.bolbilim.com) neferi olarak asıl hedefim, [tekrar edilebilir araştırma yazısı](http://bolbilim.com/2016/03/05/sabunlamadan-sisirmeden/) için makara da olsa bir rapor ortaya çıkarmak. Başlıyorum.

Önce veri bulmalı... Hah, yapay öğrenme kitabı için hazırlanmış bir sayfada buldum. Gerçek veri olduğunu da yazmışlar. Güzel. Araba fiyatları, yaşları, modelleri derken istatistikle haşır neşir olacağız. Eh istatistik dendi mi, hiç şüphesiz en iyi seçenek R.

Haydi veriyi kaydedelim.

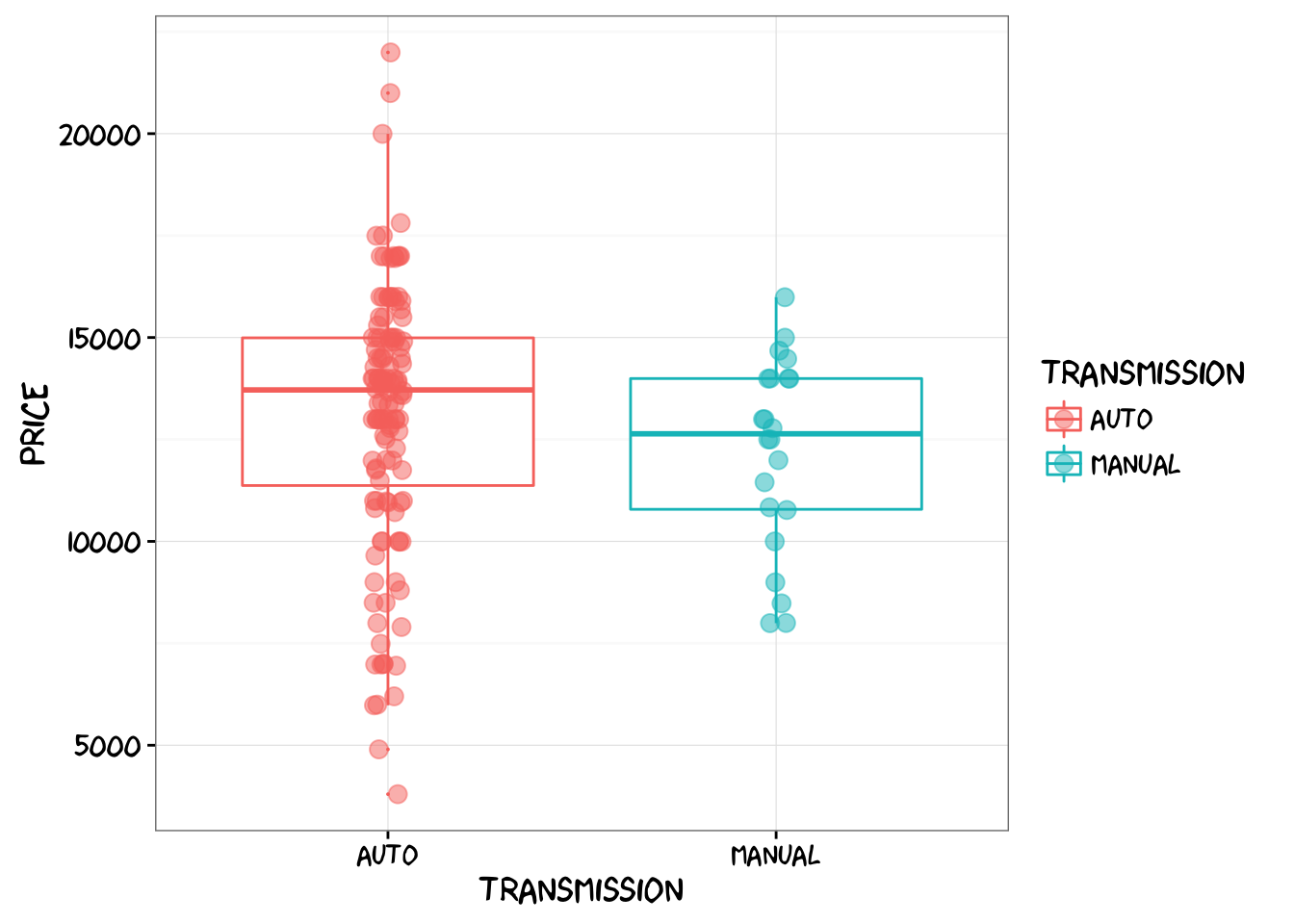
library(ggplot2)  
options(warn=-1)  
veri\_adresi <- "https://raw.githubusercontent.com/stedy/Machine-Learning-with-R-datasets/master/usedcars.csv"  
veri <- read.csv(veri\_adresi)  
veri$year <- factor(as.character(veri$year), levels=2000:2012)  
veri$color <- factor(as.character(veri$color), levels=unique(veri$color))  
cat('Toplam araba sayısı: ', nrow(veri))

## Toplam araba sayısı: 150

Çok da fazla araba yokmuş. Yine de işimizi görür.

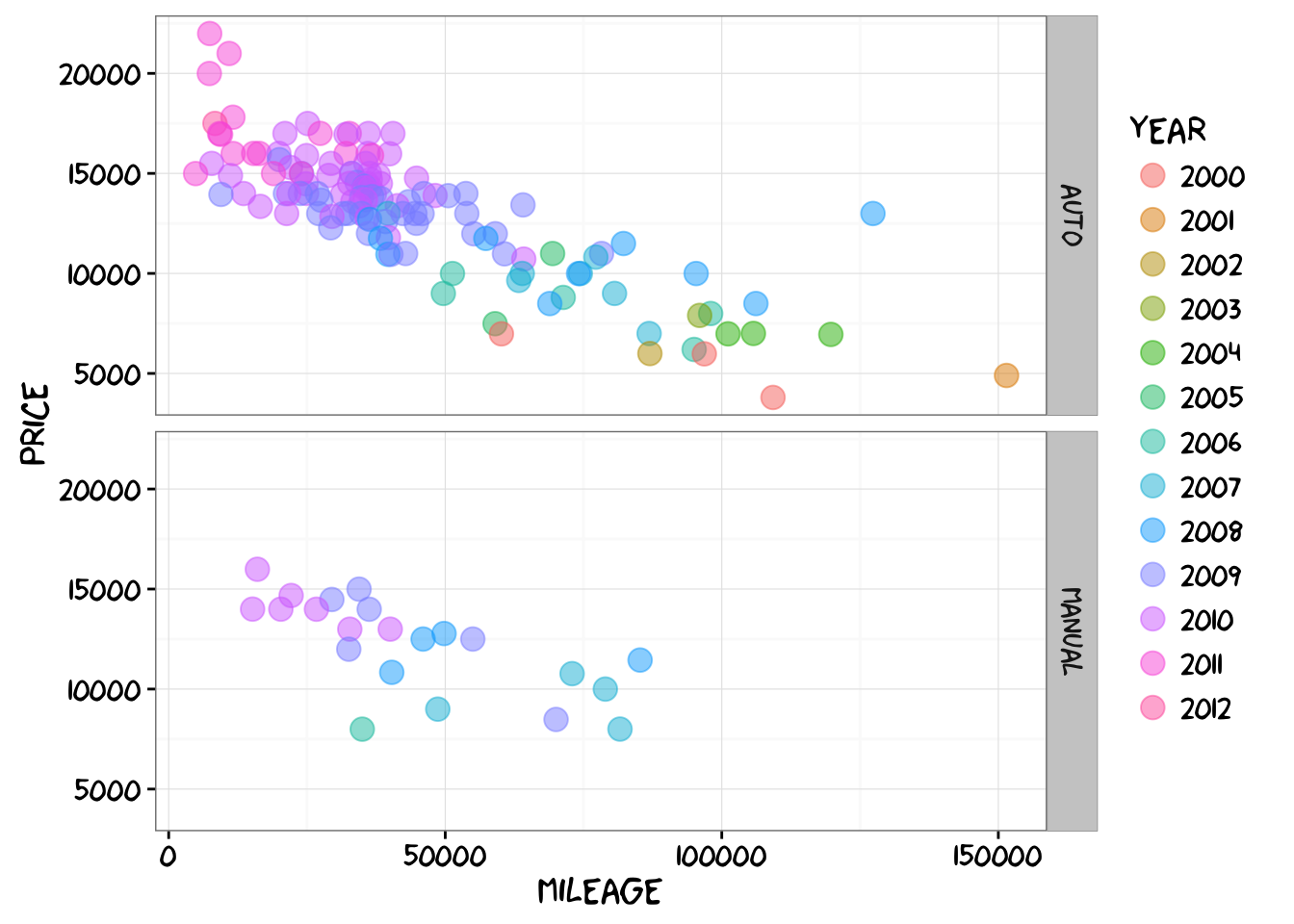
Türkiye'nin aksine, Amerika'da otomatik vites arabalar çok daha yaygın. Onun için vitesli olanlar daha ucuz olur diye bir söylenti her daim vardır. İlk ona bakalım.

ggplot(data=veri, aes(x=transmission, y=price, color=transmission)) + theme\_bw() +   
 # Bir alttaki satır, xkcd fontu yüzünden sorun çıkarabilir.   
 # Bilgisayarınızda bu font yoksa ', text=element\_text(size=16, family="xkcd")"'  
 # kısmını kaldırabilirsiniz ya da fontu değiştirebilirsiniz.  
 theme(legend.key = element\_blank(), text=element\_text(size=16, family="xkcd")) +  
 geom\_boxplot(outlier.size=0) + geom\_jitter(width=.1, size=3, alpha=0.5)



Bu grafik pek öyle demiyor. Evet ortalamada vitesli arabalar hesaplı. Ancak onlardan daha ucuz otomatik arabalar var. İyi de arabaların yaşlarını ve kaç kilometrede olduklarını göremedik. O bilgileri de ekleyelim.

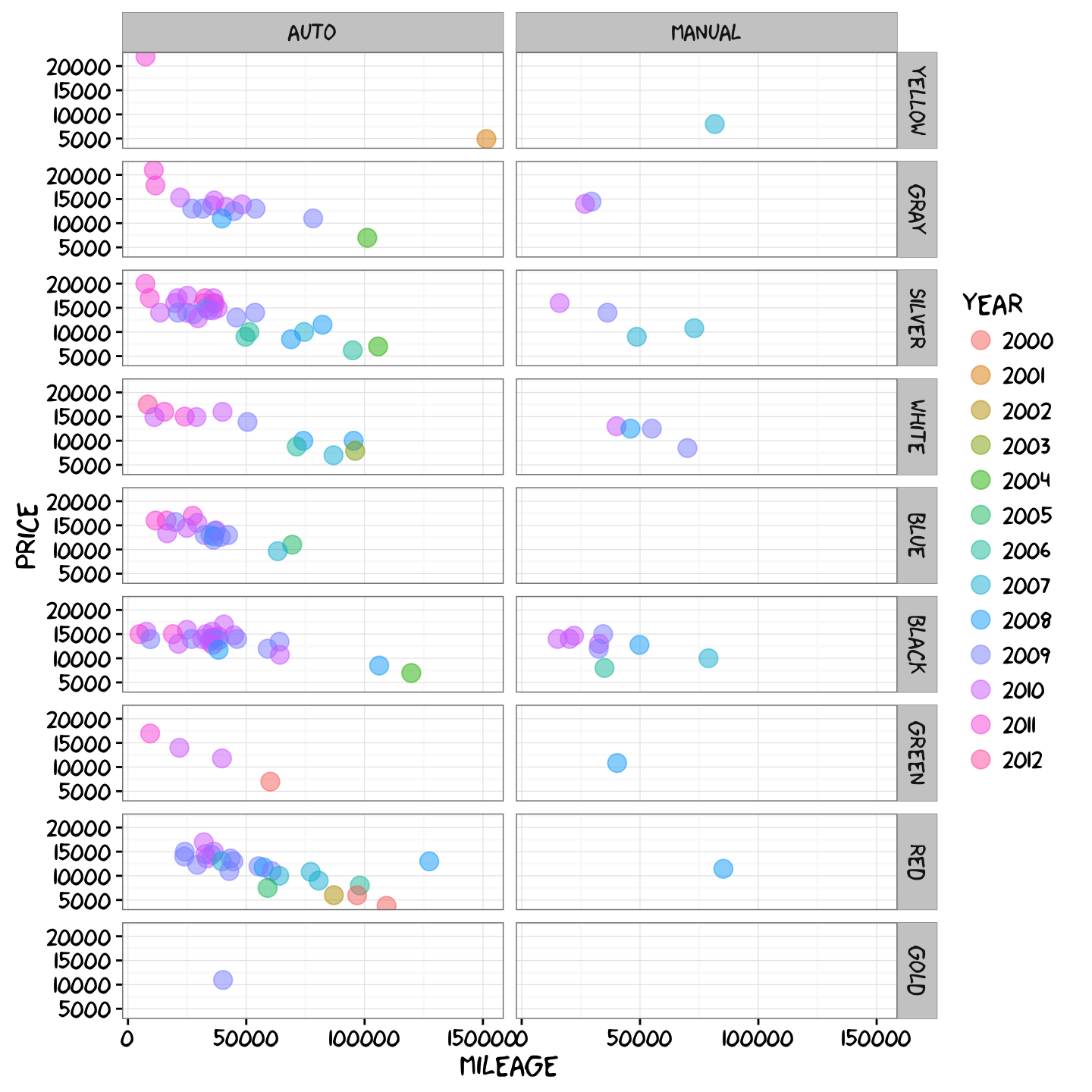
ggplot(data=veri, aes(x=mileage, y=price, color=year)) + theme\_bw() +   
 theme(legend.key = element\_blank(), text=element\_text(size=16, family="xkcd")) +  
 geom\_jitter(size=4, alpha=0.5, height = 0.5, width = 0.5) +  
 facet\_grid(transmission ~ .)



Beklenen bir grafik. Araba genç ve üzerindeki kilometre azsa, fiyatı artıyor. Şöyle bir karşılaştırma yapınca, 5-6 yaşında ve görece düşük kilometreli bir araba vitesli ise daha hesaplı oluyor.

Elimizde model verisi yok ama renk var. Sırf meraktan bir de renkleri ayıralım. Hem grafik de güzel gözüküyor.

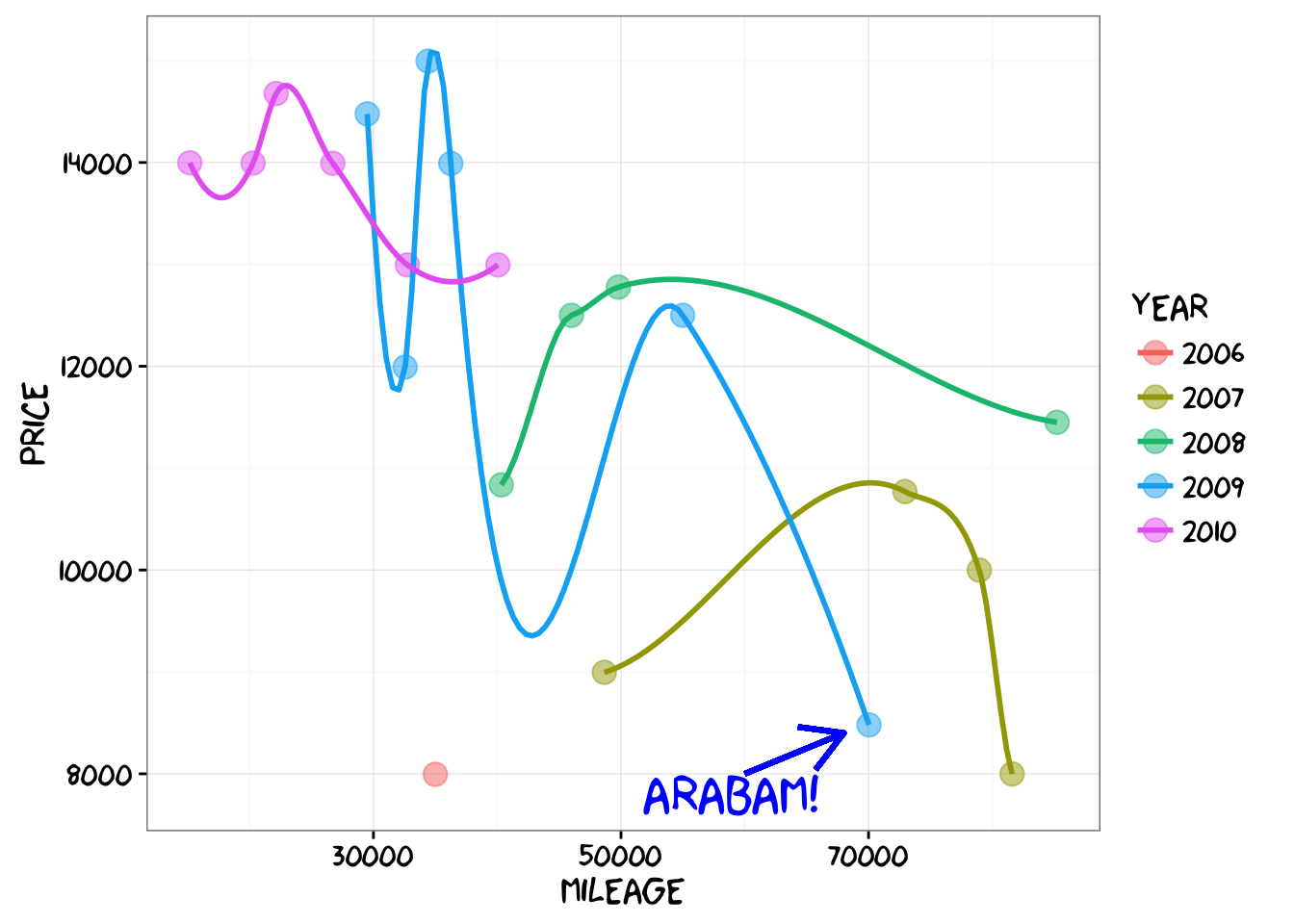
ggplot(data=veri, aes(x=mileage, y=price, color=year)) + theme\_bw() +   
 theme(legend.key = element\_blank(), text=element\_text(size=16, family="xkcd")) +  
 geom\_jitter(size=4, alpha=0.5, height = 0.5, width = 0.5) +  
 facet\_grid(color ~ transmission)



Bu grafik bize ... eee... altın renkli araba boşuna aramayın diyor. Peki. Zaten fukara öğrenci için rengin ne önemi var?

Neyse özetle, birkaç yıllık, az yorulmuş bir arabanın peşine düşerseniz, yolda kalmayacağınız ve bütçeyi tüketmeyeceğiniz bir araba için vitesli olanları düşünün. Eldeki veriye bu gözle tekrar bakınca, ben bir tane buldum. Beyaz renkli falan ama bir süre idare ederim.

ucuz <- subset(veri, transmission=="MANUAL" & price < 15000)  
ggplot(data=ucuz, aes(x=mileage, y=price, color=year)) + theme\_bw() +   
 theme(legend.key = element\_blank(), text=element\_text(size=16, family="xkcd")) +  
 geom\_point(size=4, alpha=0.5) + geom\_smooth(se=FALSE) +   
 annotate("text", x=59000, y=7800, label="ARABAM!", family="xkcd", size=8, color="blue") +  
 geom\_segment(aes(x=60000, y=8000, xend=68000, yend=8400), arrow=arrow(), size=1, color="blue")



Size gelince, yok eğer illa otomatik arabada ısrar ederseniz, elimde çok güzel, sadece ve sadece *150 bin* milde, sarı renkli bir araç var. Bir düşünün.

Kılavuzu karga olanın...