**Boxplot (Kutu Grafik)**

Boxplot, sayısal değişken görselleştirme tekniğidir.

Boxplot, histogram grafikleri gibi sayısal-sürekli değişkenleri görselleştirmek için kullanılmaktadır ve histogram ile eşit ağırlığa sahiptir diyebiliriz.

Bir sayısal değişkenin dağılımı incelenmek isteniyorsa mutlaka histogram ve boxplot kullanılmaktadır.

Veri setiyle ilk buluştuğumuz anda ilk veri setini tanımamız gerekmektedir.

Veriden değer üretmek için onu tanımamız gerekmektedir.

Veri setini tanımadan iyi-kötü olarak bu analizleri, gereksinimleri bilemeyiz

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

Bir restoranda meydana gelen, bahşişlerle ilgili yarı-yapısal veya yapısal olmayan bir veri setidir

Bunu söyleyebiliyoruz çünkü elimizdeki veri seti insan ile ilgili.

Nasa -> yapısal

Price -> yarı yapısal

Tips -> yapısal olmayan “diyebiliriz!”

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

Bu veri setinde ilk odaklanmak istediğimiz değişken “yemeğin toplam tutarı” olabilir

İkinci olarak “masada oturan kişi sayısı” olabilir

Bunlar bahşiş verme – vermeme durumlarında etkili olabilir yorumunu yapıyoruz bir girişte

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

Total\_bill (toplam fatura) ortalama 19 dolar diyebiliriz

Min değeri 3 dolar max ise 50 dolar

İçerisindeki dağılım yani çeyrekler gayet eşit güzel görünüyor (13 – 17 - 24)

Ortalama bırakılan bahşiş 3 dolar

Ve min 1 dolar bahşiş bırakılmış (vahit hoca min bıraklan bahsısın 0 olmasını beklermiş)

Masada oturan ortalama 2.5 kişi varmış – min1 max6 kişi yemek yemiş masalarda

Vahit hoca diyor ki

Masada oturan kişilerin sayısı fazla olursa? Bu bahşişin artmasına sebep olur mu?

Daha çok kişi olursa daha çok bahşiş gelir mi? (Bahşiş tutarı ve Masada oturan kişi sayısı arasında ilişki !)

Veya

Yemek yiyen masada sigara içen ve içmeyen kişiler durumunda bahşişe yansıyan tutar arasında fark yaratıyor mu?

Veya kadın olması ?

Bu tür yorumlar gerçekleşiyor!

KATEGORİK DEGİSKENLERİN SINIFLARINA

VE BU SINIFLARIN FREKANSLARINA BAKALIM

Frekans = Tekrar sayısı = kaç adet olduğu

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

Value\_counts = değerlerin sayısı! Kaç adet var ?