**İstatistiksel Düşünce Modelleri**

İstatistiksel Düşünce Modelleri: Veri Okuryazarlığından, veri analitiğine giden yolu **modelleyen** yol göstericilerdir.

Bu modeller sayesinde veriye analitik bir gözle bakabileceğiz.

**İstatistiksel Düşünce Modelleri**

metin, iş kartı, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Ben- Zvi ve Friedlander – Jones ve diğerleri daha çok okul eğitiminde kullanılan İstatistiksel Düşünce Modelleridir.

Bu yüzden analitik konusunda Mooney seçeneği bize en yakınıdır.

**Mooney Veri Modeli**

Diğer modellere kıyasla iş sektörüne daha uygun entegre edilebilmesi ve daha kolay olması nedeniyle bizim tercih ettiğimiz modeldir.

**Mooney Modeli Aşamaları**

metin, bitki, kuş içeren bir resim

Açıklama otomatik olarak oluşturuldu

1. **Verinin Tanımlanası**: Çalışmanın önemli ayrıntılarını anlamayı ve etkili bir şekilde izlemeyi sağlar.
2. **Verinin Organize edilmesi ve indirgenmesi**: Verilerin daha hızlı erişilebilmesi, işlenebilmesi ve analiz edilebilmesi için sınıflandırma sürecidir.
3. **Veri gösterimi**: verileri insan beyninin anlaması ve içgörü alması için, daha kolay hale getirmek üzere grafik gibi görsel bir bağlama çevirme uygulamasıdır
4. **Veri analizi ve yorumlama**: Temel bir ifadeyle; işlenmemiş verinin toplanması ve istatistik yöntemleri kullanarak anlamlı ve yararlı bilgiler haline getirme işlemidir.

**İstatistiksel Düşünce Düzeyleri**

1. Kişiye özgülük (Seviye 1) (Okuyamaz)
2. Geçici (Seviye 2) (Biraz okuyabilir)
3. Nicel (Seviye 3) (Okuyabilir)
4. Analitik (Seviye 4) (Okuyabilir ve Analiz Edebilir)

**Mooney Modeli Adımları**

1. **Verinin Tanımlanması**

**Örnek Senaryo**

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Sorular**

1. **Çalışmada ölçülmeye çalışılan değişkenler nelerdir?**

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

1. **Değişkenlerin Türleri Nelerdir?**

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

1. **Değişkenlerin ölçek türleri**

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

1. **Metinde dikkat çeken anomal durumlar var mı?**
2. Erkek ve Kadın oranlarının tamamı %100 değildir.
3. “Erkeklerin %72 ve Kadınların %75’i üniversite mezunudur” cümlesi hatalıdır.
4. **Veri seti yorumlama kısmı**

Gözlem birimi 2000 kişidir ve Türkiye için gayet yeterli bir sayıdır.

Eğer tarafsız bir örneklem üzerinden değerlendirme yapıldıysa bu tüm Türkiye’yi temsil edebilir.

Anketi yanıtlayanların yaş kitlesi 15-27 yaş aralığındadır. Demek ki bu şirketin hedef kitlesi genç kitledir.

Kadınların %72’si ingilizce bilmektedir kısmından çıkarım yapmak çok güçtür.

Çünkü geri kalan %28 hakkında hiçbir bilgi verilmemiştir veya Erkeklerin dil bilgisi konusunda bir bilgisi bulunmamaktadır.

Bu tür eksik tanımlamalarda karar almamak en doğrusudur. Aksi halde analizler pratikte ters sonuç verebilir.

Erkeklerin %72’si ve Kadınların %75’i Üniversite mezunudur.

Demek ki kitlenin önemli bir kısmı üniversiteden mezun olmuş gençlerdir.

Bu gayet güzel ve temiz bir bilgidir. Veri setini yorumlarken bu şekilde çıkarımlar oldukça faydalıdır.

…

1. **Verinin Organize Edilmesi ve İndirgenmesi**

Görseldeki tabloya nasıl bir düzenleme yapılmalıdır? Kısaca bahsediniz.

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

Belirli bir saat aralıklarını alarak toplulaştırma, sınıflandırma işlemi yapılabilir.

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

Sınıflandırma işleminin ardından veri setinin eksiklikleri, veri setinin ölçülebilirliği gibi konular çok daha net şekilde görülebilmektedir.

Bu yüzden verileri analiz etmeden önce veri setini sınıflandırmak oldukça önemli bir adımdır.

Eksik verileri, gürültülü verileri bu sayede giderebilir ve analize devam edebilirsiniz. Bu işlemleri yapmadan analiz adımına geçmek yanıltıcı sonuçlar doğuracaktır.

1. Verinin Gösterimi

metin, ekran görüntüsü, yazı tipi, diyagram içeren bir resim

Açıklama otomatik olarak oluşturuldu

Grafik2 doğru olarak verilmiştir.

Grafik gösterimlerinde x ve y eksenlerinde ölçekler eldeki verinin sunumunda suistimale en açık olan yerlerdir.

Başlangıç noktaları sıfır olmak zorundadır.

Grafik Ölçülerinin artışları Grafik2’ de daha doğru verilmiştir (+10.000 Artış)

Grafik1 için bu artış seçenekleri çok yanıltıcı grafik sonucu vermiştir! (+1000)

NOT: Veri ile ilgili bir işlem yaparken iş bilgisi (sektör bilgisi) bu gibi durumlarda çok önemlidir.

Karar verirken sektör içerisinden birisi gibi düşünmeli, bakmalıyız. İşte bir Veri Bilimcinin yorumu, analizi tam olarak burada devreye girmektedir.

1. Verinin Analiz Edilmesi ve Yorumlanması

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

Ortalamaya bakarak takımı (veri setini) işleme, karara tabii tutarsak hatalı analiz gerçekleşir.

Çünkü Standart sapmasına baktığımız zaman (ortalama etrafındaki dağılım) dağılımın geniş olduğunu görüyoruz.

Gönüllüler takımındaki yoğunluk (homojenlik) çok daha doğru bir seçimdir çünkü standart sapmaları düşüktür.

Gönüllüler takımındaki insanların güç oranları daha uygun ve uyumludur.

Diğer yandan Ünlüler takımının ortalaması yüksektir fakat standart sapmaları yani ortalamaya olan uzaklıkları da çok yüksektir.

Bu gürültülü veri olduğu anlamına gelir ve güvensizdir.