**Rapor:**

**Sigorta Veri Seti Analizi ve Öneriler**

**1. Prim Fiyatlandırması**

**1.1. Z-Score Yöntemi ile Aykırı Değerlerin Tespiti**

* Z-Score yöntemi kullanılarak (+2, -2) aralığında aykırı değerler tespit edildi.

**1.2. Yaşa Göre Prim Fiyatlandırması**

* Age sütunu belirli gruplara ayrılarak ['18-29', '30-39', '40-49', '50-59', '60+'] yaş grupları oluşturuldu.
* Her yaş grubunun ortalama prim fiyatı hesaplandı.

**1.3. Sigara Kullanımına Göre Prim Fiyatlandırması**

* 'smoker' sütununa göre 'yes' veya 'no' cevabı veren kişilerin ortalamaları hesaplandı.

**2. Risk Profil Analizi**

**2.1. Demografik Gruplara Göre Risk Profili**

* Yas, cinsiyet, BMI ve smokerlık durumuna göre risk profilleri analiz edildi.
* Örneğin, cinsiyete göre male ve female ortalamaları ortaya çıkarıldı.
* BMI değerlerinin yüksek ve düşük olanları ele alınarak analizler yapıldı.

**3. Dolandırıcılık Tespiti**

* Veri setindeki tutarsızlıklar ve anormallikler izlenerek potansiyel dolandırıcılık vakaları tespit edildi.
* Sigara içmeyen kişilerin, genellikle sigara içenlere göre daha düşük sigorta primleri ödemeleri beklenir. Ancak, bu listedeki bazı sigara içmeyen kişilerin oldukça yüksek prim ödeme tutarları dikkat çekicidir.
* Dolandırıcılık Olabilir: Yüksek prim ödemeleri, bu kişilerin sigara içmeyenler olarak etiketlendiği halde, sigara içenlerle benzer veya daha yüksek prim ödeme durumlarına işaret edebilir. Bu durum, potansiyel dolandırıcılık veya veri girişi hatalarının bir göstergesi olabilir.
* Detaylı İnceleme Gerekebilir: Bu örneklerin detaylı bir incelemeye tabi tutulması gerekebilir. Özellikle, sigara içmeyen bir kişinin neden yüksek bir prim ödediği belirlenmeli ve bu durumun sistematik bir sorun olup olmadığı değerlendirilmelidir.

**4. Düşük Sigorta Primleri ve İleri Analiz**

* Sigara içmeyen kişilerin primleri genellikle düşük olmalıdır. Ancak, bazı durumlarda anormal yüksek prim ödemeleri tespit edildi.
* Bu durumlar detaylı bir inceleme gerektirir ve makine öğrenimi modelleri kullanılarak daha kapsamlı analizler yapılabilir.

**5. Pazarlama Stratejileri**

* Analiz sonuçlarına dayanarak belirli demografik gruplara özel pazarlama stratejileri geliştirilebilir.
* Özelleştirilmiş hizmetler sunarak müşteri çekme stratejileri optimize edilebilir.

**6. Önerilen Veri Analizi ve Filtrasyonlar**

**6.1. Sigara Kullanımı Analizi**

* Operatörü "esit degildir" kullanarak "smoker" sütununda "yes" olan verileri filtreleyerek sigara içen müşterilerin sigorta maliyetlerini inceleyebiliriz.

**6.2. BMI Kategorileri Analizi**

* BMI değerlerine göre belirli aralıkları filtreleyerek BMI kategorilerini oluşturabilir ve bu kategorilerdeki müşterilerin maliyetlerini karşılaştırabiliriz.

**6.3. Yaş Grupları Analizi**

* Operatörleri kullanarak belirli yaş aralıklarına göre filtreleme yapabilir ve farklı yaş gruplarındaki müşterilerin maliyetlerini inceleyebiliriz.

**4. Cinsiyet Analizi:**

* "sex" sütununu kullanarak cinsiyet bazlı filtreleme yapılacak ve cinsiyetin maliyetlere etkisi değerlendirilecektir.

**5. Bölgelere Göre Analiz:**

* Operatörleri kullanarak belirli bölgelerdeki müşterileri filtreleyebilir ve coğrafi bölgeler arasındaki maliyet farklarını detaylı bir şekilde inceleyebiliriz.

**6. Çocuk Sayısının Etkisi:**

* Operatörleri kullanarak belirli çocuk sayılarına göre filtreleme yapabilir ve çocuk sayısının maliyetler üzerindeki etkisi değerlendirilecektir.