**Izveštaj za Investitora i moguća poboljšanja**

**Tema: EDA Mesečne Prosečne Potrošnje Domaćinstava**

**Uvod**

U okviru našeg projekta analizirali smo dataset koji se odnosi na mesečnu prosečnu potrošnju domaćinstava, uzet je period od 2006 do 2022. Cilj analize bio je da se bolje razumeju obrasci potrošnje, odnosno da se identifikuje na šta se najviše troši i kako se to kretalo sa godinama. Takođe je razvijen model za predviđanje budućih trendova potrošnje.

**Ključni Nalazi. Analiza i Čišćenje Podataka**

Obrađeni su podaci: proverio sam da li postoje duplikati i kakva je struktura podataka kao i osnove informacije o tabelam i kolonama što je bitno za dalju analizu.

Struktura podataka: Dataset sadrži različite tipove podataka, uključujući vremenske serije i kategorije potrošnje, što smo koristili za detaljnu analizu.

**Vizualizacije**

Distribucija i trendovi: Kreirane su raznovrsne vizualizacije koje prikazuju distribuciju i trendove potrošnje, omogućavajući lakše razumevanje podataka.

Vizualizacije pružaju jasne uvide u ponašanje potrošača, identifikovanje sezonskih uzoraka, i varijacije u potrošnji između različitih grupa.

**Model Mašinskog Učenja**

Implementacija modela: Upotrebljen je model Random Forest za prognoziranje potrošnje, koji pokazuje visok nivo tačnosti.

Rezultati modela: Model je efikasno identifikovao važne faktore koji utiču na potrošnju, što može pomoći u donošenju informisanih odluka.

**Preporuke za Dalje Istraživanje. Prikupljanje Dodatnih Podataka**

Dodavanje dodatnih varijabli kao što su dodatni ekonomski pokazatelji, demografski podaci, poređenje sa istim trendovima u regiji, mogla bi se unapredila analiza. Takođe pošto su kao pokazatelj za plate korišćeni samo prosek na državnom nivou, detaljniji izveštaj po opštinama ili regijama bi poboljšao analizu.

**Dublja Analiza Podataka**

Neke dodatne tehnike za analize i modele, korelacijska analiza i uzročna inferencija, za bolje razumevanje kako različiti faktori utiču na potrošnju.

Implementacija specifičnih modele za analizu vremenskih serija da bi dobili dublje uvide u sezonske obrasce i dugoročne trendove.

**Segmentacija Kupaca**

Upotreba metode klasterovanja za identifikaciju različitih segmenata potrošača, što može pomoći u ciljanom marketingu i optimizaciji ponude. (starost, demografski podaci, regija)

**Prediktivna Analitika**

Istraživanje naprednijeg modela prediktivne analitike ili tehnike dubokog učenja za preciznije predviđanje budućih trendova potrošnje.

**Zaključak**

Ova analiza pruža temeljan uvid u obrasce potrošnje domaćinstava, identifikujući ključne faktore koji utiču na nju. Uvođenjem dodatnih podataka i primenom naprednijih analitičkih metoda, možemo dalje proširiti naše razumevanje i unaprediti sposobnost predviđanja budućih trendova. Ovi uvidi i predviđanja su ključni za donošenje informisanih odluka i prilagođavanje strategija na tržištu.

**Rizici i Izazovi**

Postoji potencijalni rizik od prekomerne pristrasnosti u modelima ukoliko se koriste neadekvatni ili pristrasni podaci. Kompleksnost modela i interpretacija rezultata može biti izazov za osobe koje nisu upoznate s analitičkim tehnikama. Ovaj izveštaj predstavlja jasan plan za dalje korake i razvoj projekta, uz naglasak na inovativne pristupe analizi podataka i prediktivnoj analitici.