



گزارش تکلیف اول درس الگوریتم های علوم داده (بخش دوم)

نام و نام خانوادگی: فاطمه ترودی

شماره دانشجویی: ۴۰۳۴۲۲۰۴۸

نام استاد: دکتر سعیدرضا خردپیشه

نیمسال دوم ۱۴۰۳-۰۴

فهرست مطالب

۱. مقدمه.....	۳
۲. تحلیل اکتشافی داده‌ها (EDA).....	۳
۲.۱. بررسی داده های گم‌شده.....	۳
۲.۲. تحلیل تک‌متغیره.....	۵
۲.۳. تحلیل دو متغیره.....	۱۱
۳. آزمون های فرض.....	۱۸
۳.۱. آزمون فرض اول.....	۱۸
۳.۲. آزمون فرض دوم.....	۱۹
۳.۳. آزمون فرض سوم.....	۱۹
۳.۴. آزمون فرض چهارم.....	۱۹
۳.۵. آزمون فرض پنجم.....	۲۰

۱. مقدمه

درک رفتار مشتری برای کسب و کارهایی که هدفشان بهبود استراتژی های بازاریابی، افزایش تعامل با مشتری و افزایش فروش است، بسیار مهم است. در این تمرین، تجزیه و تحلیل شخصیت مشتری را انجام می دهیم تا عادات خرید مشتری، الگوهای هزینه و ... را در را بررسی کنیم. تجزیه و تحلیل ما بر اساس مجموعه داده ای است که حاوی اطلاعات مشتری مانند سن، درآمد، وضعیت تاهل، سطح تحصیلات و رفتار هزینه در انواع مختلف محصول است. با استفاده از تجزیه و تحلیل داده های اکتشافی (EDA) و تکنیک های آماری، هدف ما کشف نتایج کلیدی در مورد رفتار های مشتری است. در این تمرین موارد زیر مورد بررسی قرار می گیرند:

- تحلیل اکتشافی داده ها برای درک بهتر رفتار های مشتری و عوامل تاثیرگذار در خرید آنها
- انجام آزمون های فرض آماری برای پاسخ به سوالات کلیدی مهم.

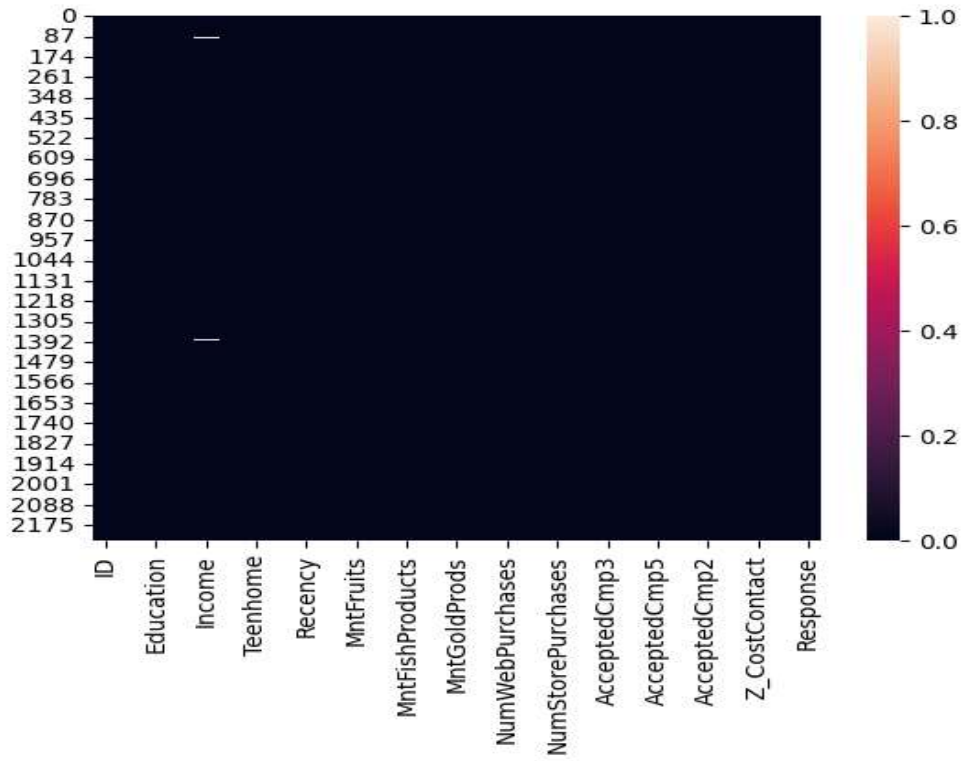
۲. تحلیل اکتشافی داده ها (EDA)

خلاصه مجموعه داده ها نکته های کلیدی را در مورد ساختار و کامل بودن آن ارائه می دهد. این مجموعه داده شامل ۲۲۴۰ سطر در ۲۹ ستون است.

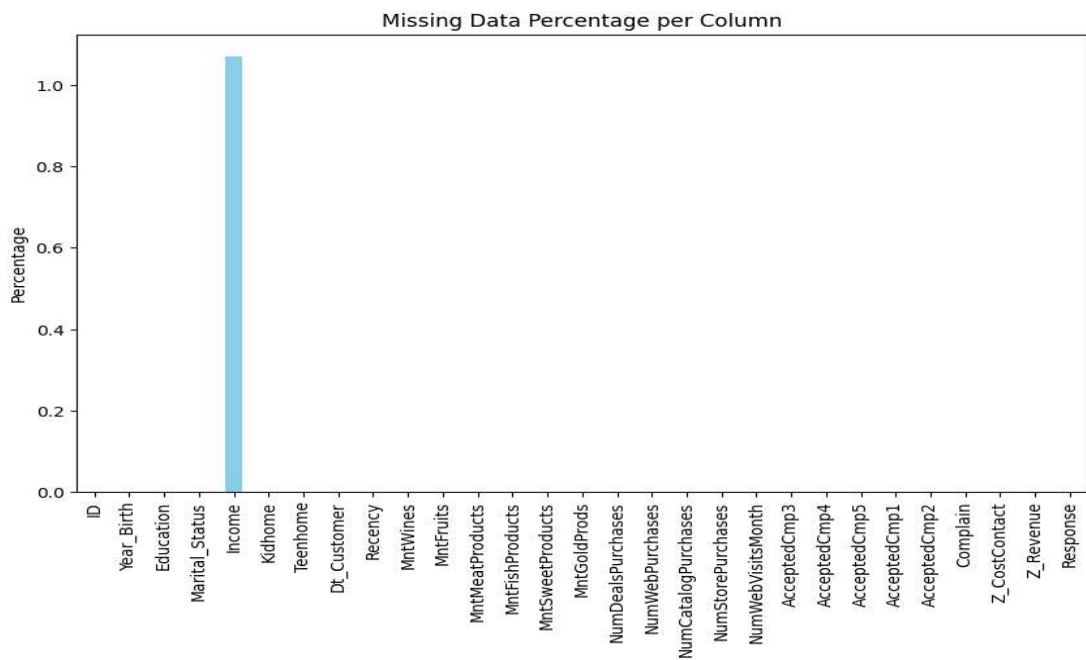
این مجموعه داده شامل ۲۹ متغیر از جمله سال تولد، سطح تحصیلات، وضعیت تاهل، درآمد، هزینه های مختلف بر اساس گروه های محصولات و ... می باشد که ما در این بخش با استفاده از رسم نمودار های مفید و کاربردی به تجزیه و تحلیل آنها پرداخته ایم.

۲/۱. بررسی داده های گم شده

در این بخش مقادیر گم شده در هر متغیر بررسی شده است و در نمودار میله ای زیر تعداد مقادیر گم شده را مشاهده می کنیم.



شکل ۱- نمودار حرارتی مقادیر گم شده



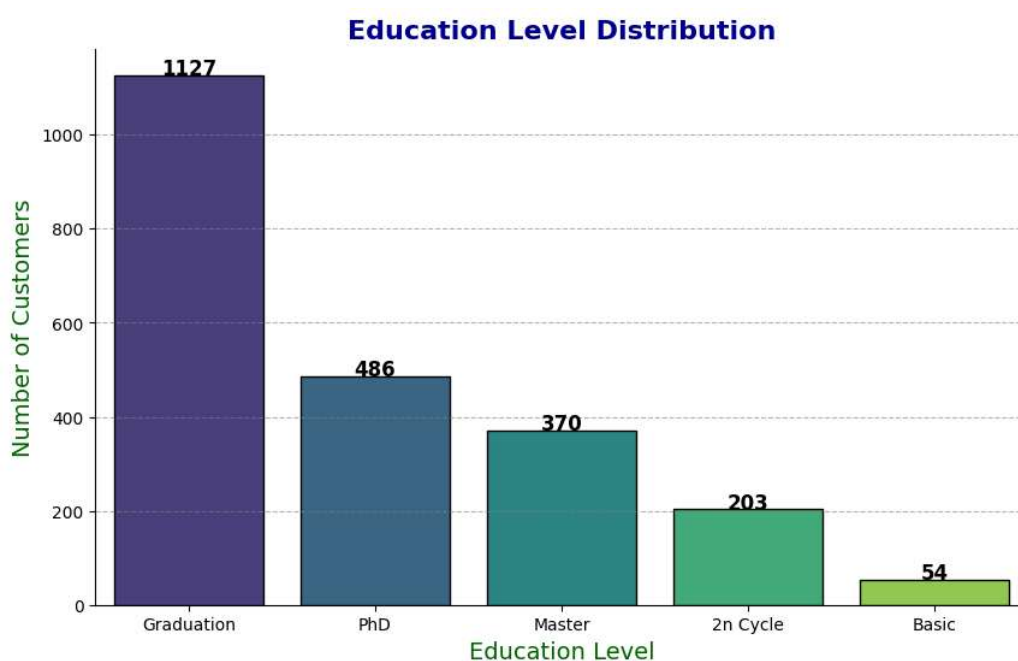
شکل ۲- نمودار میله ای درصد مقادیر گم شده متغیرها

مشاهده می‌شود که فقط متغیر Income دارای ۲۴ مقدار گم‌شده می‌باشد که حدود ۱ درصد از کل مجموعه داده است. ما برای حل این مورد، اقدام به پر کردن مقادیر گم‌شده با استفاده از میانگین متغیر Income کردیم.

۲/۲. تحلیل تک‌متغیره

بررسی توزیع متغیر Education

نمودار میله ای برای بررسی توزیع متغیر سطح تحصیلات مشتریان به صورت زیر می‌باشد:



شکل ۳- نمودار میله ای توزیع متغیر Education

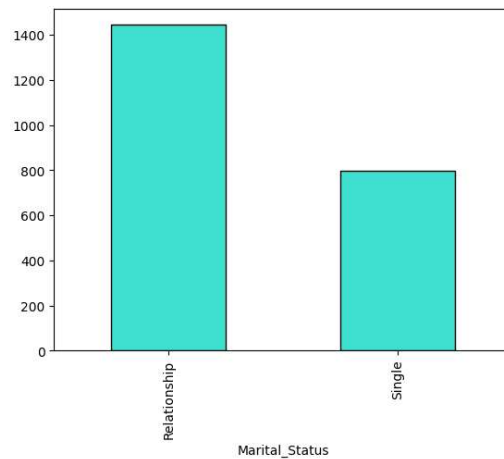
مشاهده می‌شود که ۹۷/۵۸ درصد مشتریان فارغ التحصیل و دارای مدرک تحصیلی معتبر می‌باشند و تنها تعداد بسیار کمی از مشتریان (۲/۴۱ درصد) دارای مدرک تحصیلی نمی‌باشند که نشان دهنده تحصیل کرده بودن اکثر مشتریان این فروشگاه است. برای مثال ۴۸۶ نفر از ۲۲۴۰ مشتری دارای مدرک دکتری و ۳۷۰ نفر دارای مدرک کارشناسی ارشد می‌باشند که خود به تنهایی نیز می‌تواند این موضوع را بیان کند.

بررسی متغیر وضعیت تاهل (Marital_Status)

برای بررسی وضعیت تاهل ما مشتریان را به دو گروه کسانی که در رابطه ای هستند و کسانی که در رابطه ای نیستند تقسیم کردیم و تحلیل را بر اساس این دو گروه انجام دادیم.

نمودار میله ای برای بررسی توزیع متغیر وضعیت تاهل بر اساس دو گروه ذکر شده به صورت زیر می باشد:

Frequency Of Each Category in the Marital_Status Variable

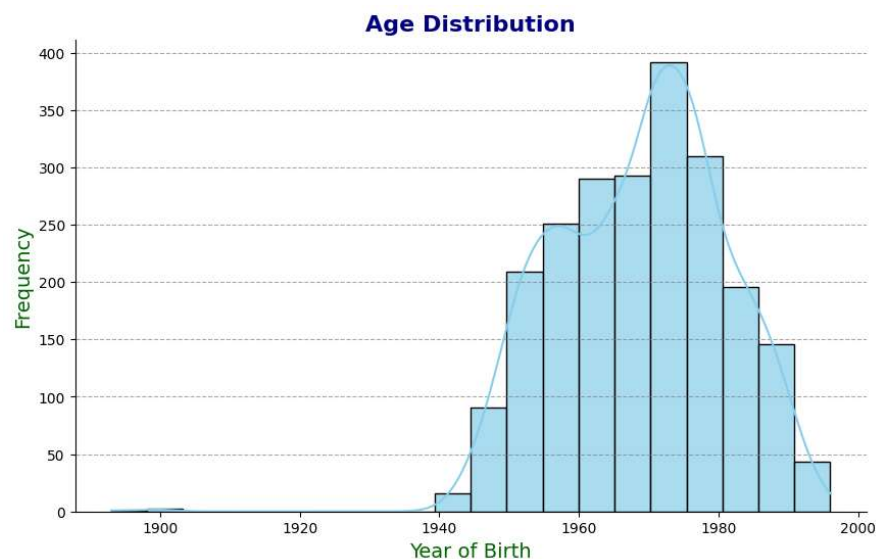


شکل ۴- نمودار میله ای توزیع متغیر بررسی توزیع متغیر وضعیت تاهل بر اساس دو گروه

مشاهده می شود که ۶۴/۴۶ درصد از مشتریان در مجموعه داده در رابطه هستند و ۳۵/۵۳ درصد از مشتریان در مجموعه داده مجرد هستند.

بررسی متغیر سال تولد و سن

برای بررسی سال تولد و سن مشتریان نمودار هیستوگرام به صورت زیر رسم شده است:

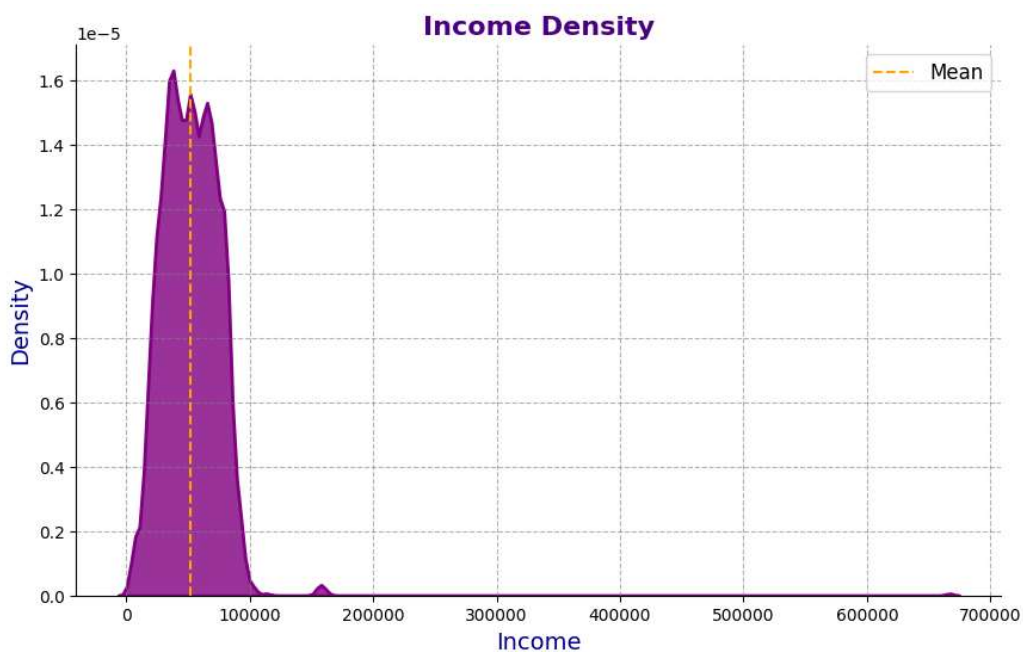


شکل ۵- نمودار هیستوگرام توزیع متغیر سال تولد

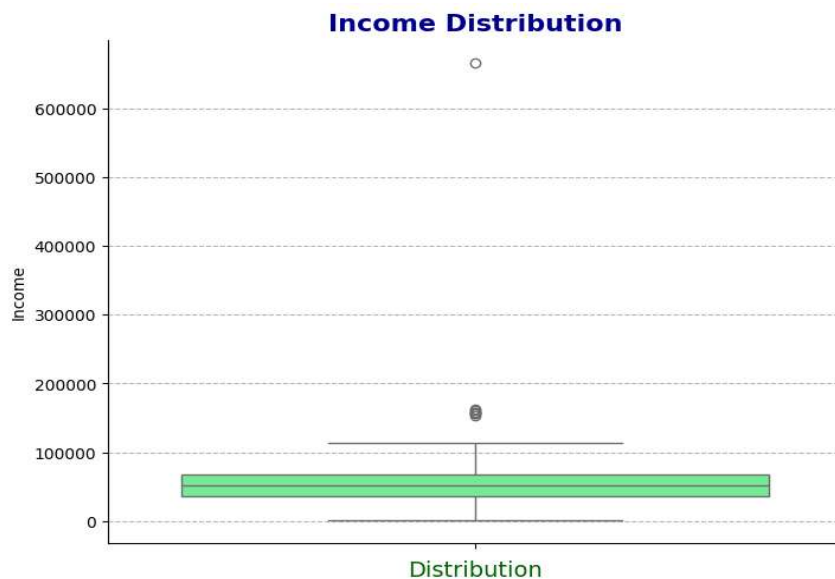
مشاهده می‌شود که بیشتر مشتریان دارای سن حدود ۵۰ سال می‌باشند. همچنین تعداد کمی از مشتریان این فروشگاه دارای سن بسیار پایین یا بسیار بالا هستند.

بررسی متغیر درآمد

برای بررسی درآمد مشتریان نمودار KDE و همینطور نمودار جعبه ای به صورت زیر رسم شده اند:



شکل ۶- نمودار KDE چگالی متغیر درآمد



شکل ۷- نمودار جعبه ای توزیع متغیر درآمد

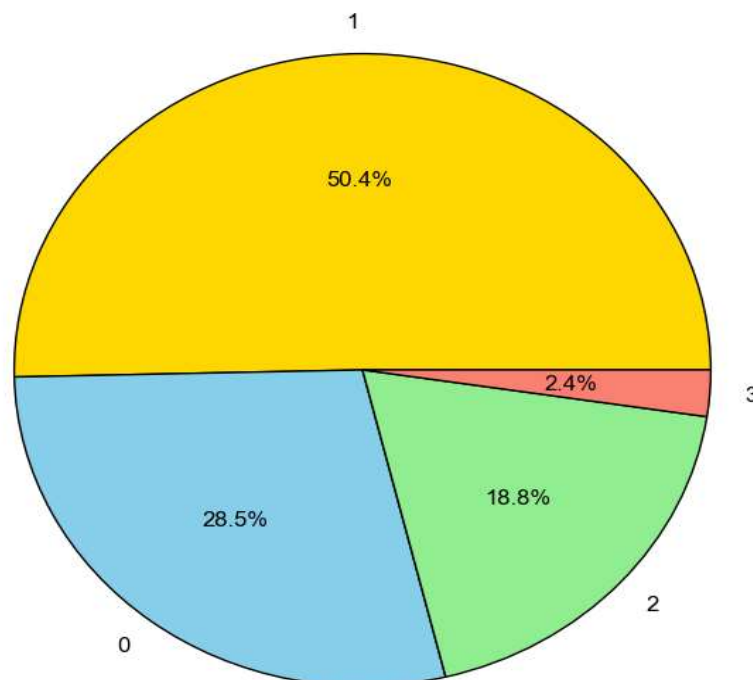
مشاهده می‌شود که میانگین درآمد مشتریان برابر ۵۲۲۴۷ دلار در سال است. همچنین کمترین و بیشترین درآمد به ترتیب برابر ۱۷۳۰ و ۶۶۶۶۶۶ دلار در سال می‌باشند. توزیع متغیر درآمد دارای چولگی به سمت راست می‌باشد و این نشان دهنده تفاوت در درآمد های مشتریان این فروشگاه است که می‌تواند بر قدرت خرید آنها تاثیرگذار باشد. توزیع درآمد مشتریان نیز دارای چند مقدار دور افتاده است که نشان دهنده وجود افرادی با درآمد های بالاتر نسبت به سایر مشتریان است.

بررسی متغیرهای Teenhome و Kidhome

برای بررسی تعداد فرزندان مشتریان یک متغیر جدید به نام Kids تعریف کردیم که حاصل از جمع دو متغیر Teenhome و Kidhome است که برابر با مجموع فرزندان یک مشتری می‌باشد.

نمودار دایره ای فراوانی تعداد فرزندان مشتریان به صورت زیر می‌باشد:

Frequency Of Each Category in the Kids Variable

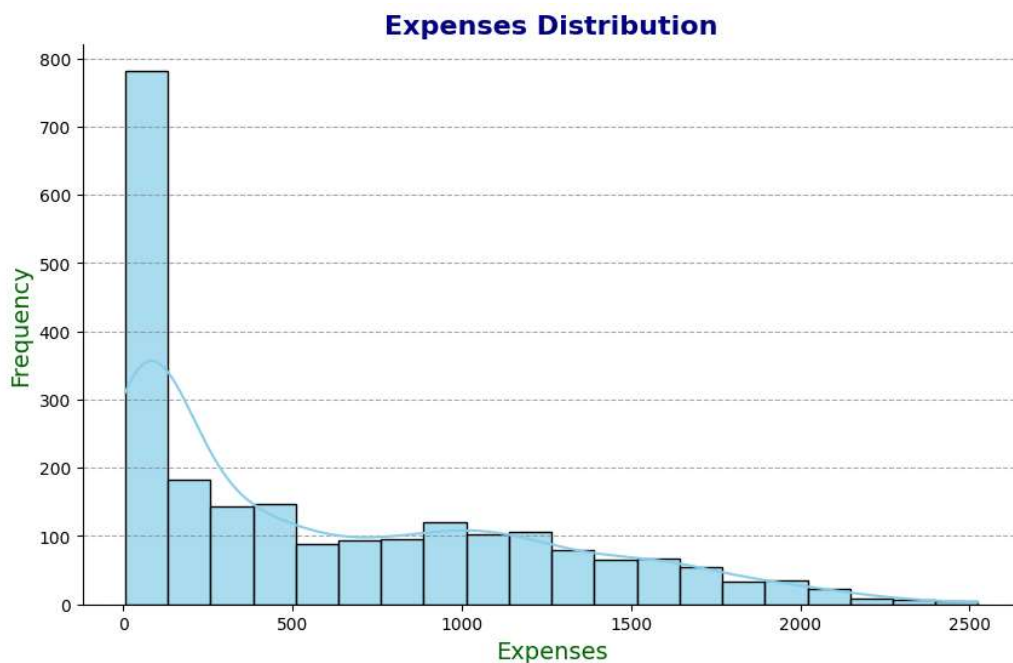


شکل ۸- نمودار دایره ای فراوانی تعداد فرزندان مشتریان

۵۰/۴ درصد از مشتریان در مجموعه داده ۱ فرزند دارند. ۲۸/۵ از مشتریان در مجموعه داده بچه ندارند. ۱۸/۸ درصد از مشتریان در مجموعه داده ۲ فرزند دارند. ۲/۴ درصد از مشتریان در مجموعه داده ۳ فرزند دارند.

بررسی هزینه مشتریان

برای بررسی هزینه مشتریان یک متغیر جدید به نام Expenses تعریف کردیم که حاصل از جمع متغیرهای MntGoldProds و MntSweetProducts، MntFishProducts، MntMeatProducts، MntWines می‌باشد که برابر با کل هزینه های مشتری حاصل از خرید نوشیدنی، محصولات گوشتی، ماهی و ... است. نمودار هیستوگرام و جعبه ای توزیع متغیر هزینه ها به صورت زیر رسم شده است:



شکل ۹- نمودار هیستوگرام توزیع متغیر Expenses



شکل ۱۰- نمودار جعبه ای توزیع متغیر Expenses

توزیع متغیر هزینه (Expenses) دارای چولگی به سمت راست است که نشان دهنده نرمال نبودن و تفاوت چشمگیر در توان خرید مشتریان است. متغیر Expenses نیز دارای چند مقدار دور افتاده است که مشتریان با بیشترین مقدار خرید از فروشگاه را نشان می‌دهد.

همچنین میانگین هزینه مشتریان برابر ۶۰۵ دلار است و نیز کمترین و بیشترین هزینه به ترتیب برابر ۵ و ۲۵۲۵ دلار می‌باشند.

بررسی خریدهای آنلاین و خریدهای فروشگاه

برای بررسی این مورد، یک متغیر جدید به نام Total_Online_Purchases تعریف کردیم که حاصل از جمع متغیرهای NumWebPurchases و NumCatalogPurchases می‌باشد و برابر تعداد خریدهای آنلاین مشتریان است. همچنین یک متغیر جدید به نام Total_Store_Purchases تعریف کردیم که برابر تعداد خریدهای فروشگاه مشتریان است.

نمودار میله‌ای فراوانی تعداد خریدهای آنلاین و فروشگاه مشتریان به صورت زیر رسم شده است:



شکل ۱۱- نمودار میله‌ای تعداد خریدهای آنلاین و فروشگاه مشتریان

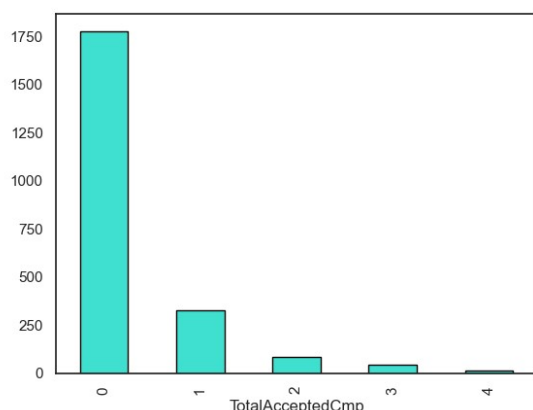
مشاهده می‌کنیم که تعداد خریدهای آنلاین بیشتر از خریدهای در فروشگاه توسط مشتریان بوده است و از کل خریدها تقریباً ۱۴۵۰۰ آنها آنلاین بوده است.

بررسی پذیرفتن پیشنهاد در کمپین ها توسط مشتریان

برای بررسی این مورد، یک متغیر جدید به نام TotalAcceptedCmp تعریف کردیم که حاصل از جمع متغیر های AcceptedCmp1، AcceptedCmp2، AcceptedCmp3، AcceptedCmp4 و AcceptedCmp5 می باشد و برابر تعداد مشتریانی است که در هر کمپین پیشنهاد فروشگاه را پذیرفته اند.

نمودار میله ای فراوانی این مورد به صورت زیر رسم شده است:

Frequency Of Each Category in the TotalAcceptedCmp Variable



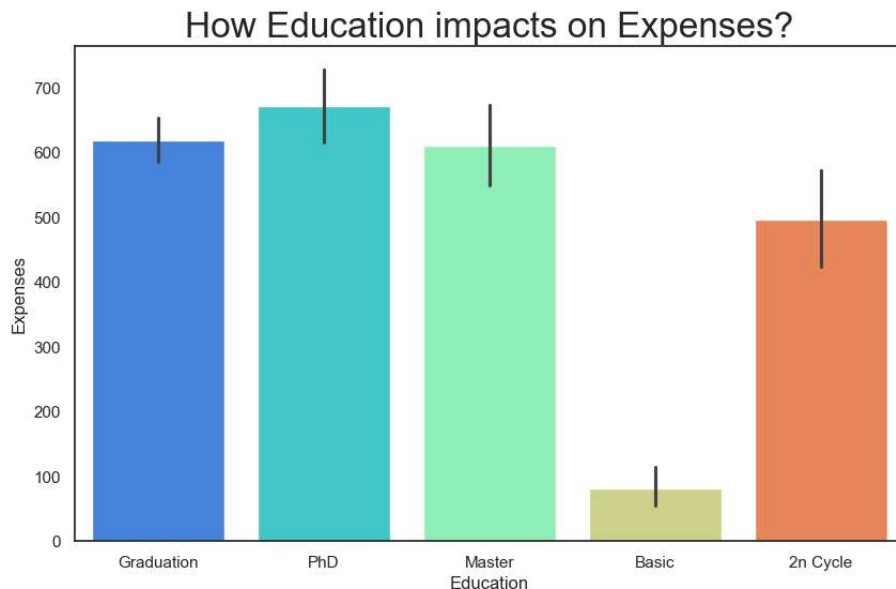
شکل ۱۲- نمودار میله ای تعداد پذیرفتن پیشنهاد فروشگاه توسط مشتریان

۷۹.۳۳ درصد از مشتریان پیشنهاد را در کمپین شماره یک پذیرفته اند. ۱۴.۵ درصد از مشتریان پیشنهاد را در کمپین شماره دو پذیرفته اند. ۳.۷ درصد از مشتریان پیشنهاد را در کمپین شماره سه پذیرفته اند. ۱.۹۶ درصد از مشتریان پیشنهاد را در کمپین شماره چهار پذیرفته اند. ۰.۴۹ درصد از مشتریان پیشنهاد را در کمپین شماره ۵ پذیرفته اند.

۲/۳. تحلیل دومتغیره

بررسی تاثیر سطح تحصیلات بر هزینه ها

نمودار میله ای برای بررسی تاثیر سطح تحصیلات بر هزینه ها به صورت زیر رسم شده است:

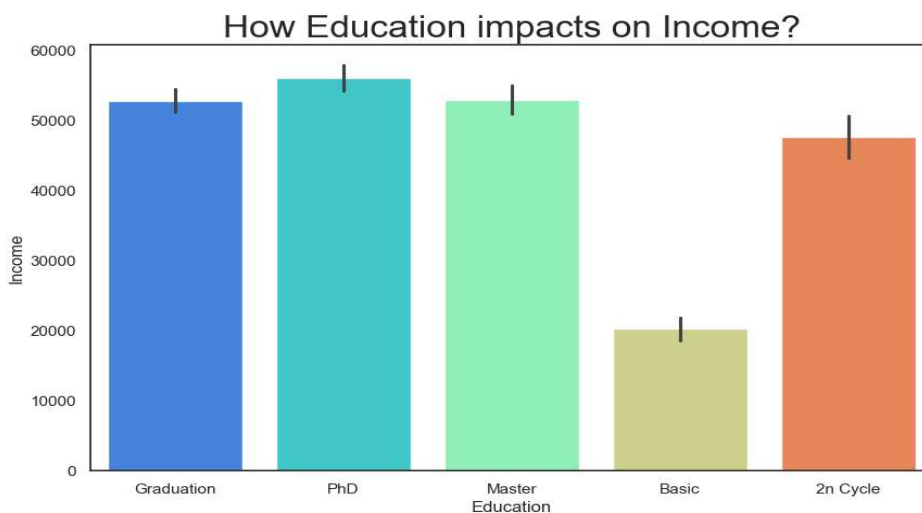


شکل ۱۳- نمودار میله ای تاثیر سطح تحصیلات بر هزینه ها

مشاهده می شود که مشتریانی که دارای تحصیلات بالاتری هستند خرید بیشتری انجام داده اند که می تواند ناشی از بالاتر بودن در آمد این افراد باشد. افراد دارای مدرک دکتری حدود ۶۵۰ دلار هزینه کرده اند که در بین تمام سطح های تحصیلات دارای بیشترین توانایی در خرید هستند. همچنین مشاهده می شود افرادی که دارای سطح پایینی از تحصیلات می باشند، توانایی خرید بسیار پایین تری حدود کمتر از ۱۰۰ دلار دارند.

بررسی تاثیر سطح تحصیلات بر درآمد ها

نمودار میله ای برای بررسی تاثیر سطح تحصیلات بر درآمد ها به صورت زیر رسم شده است:

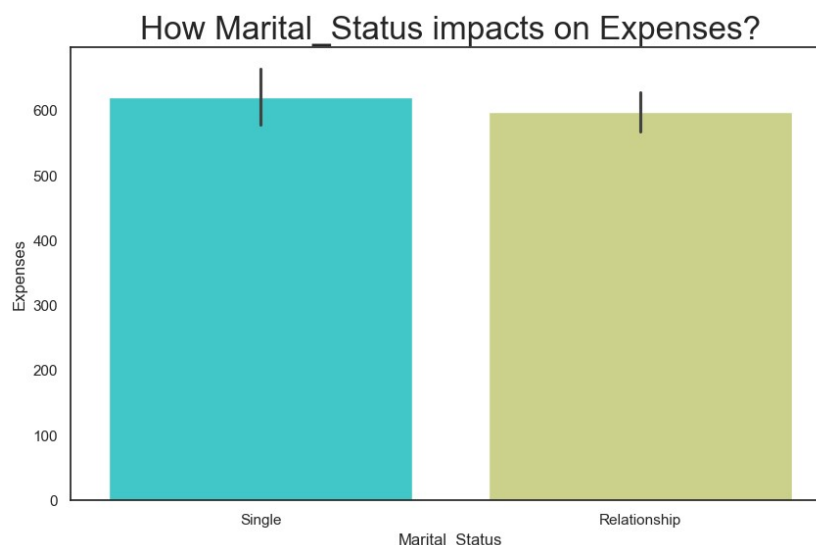


شکل ۱۴- نمودار میله ای تاثیر سطح تحصیلات بر درآمد ها

مشاهده می‌شود که مشتریانی که دارای تحصیلات بالاتری هستند درآمد بالاتری دارند انجام داده‌اند که کاملاً هم علت این موضوع واضح می‌باشد. افراد دارای مدرک دکتری حدود ۵۵۰۰۰ دلار در سال درآمد دارند که در بین تمام سطح‌های تحصیلات دارای بیشترین درآمد است. همچنین مشاهده می‌شود افرادی که دارای سطح پایینی از تحصیلات می‌باشند، توانایی کمتری دارند.

بررسی تاثیر وضعیت تاهل بر هزینه‌ها

نمودار میله‌ای برای بررسی تاثیر وضعیت تاهل بر هزینه‌ها به صورت زیر رسم شده است:

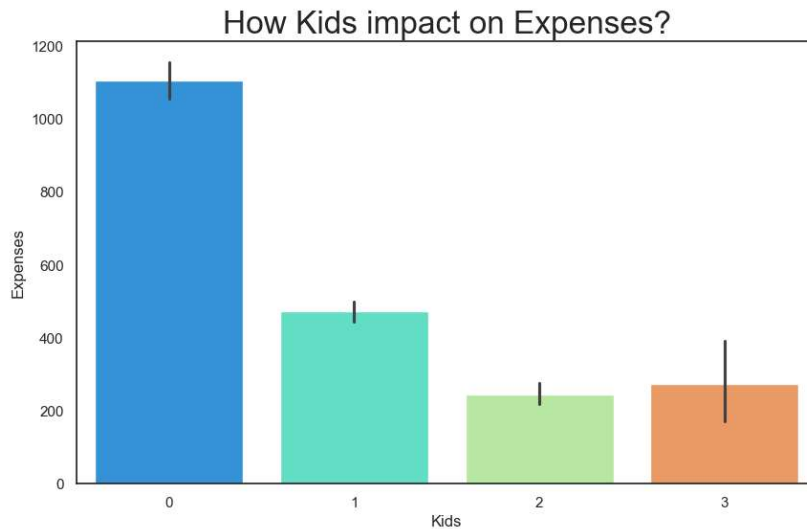


شکل ۱۵- نمودار میله‌ای تاثیر وضعیت تاهل بر هزینه‌ها

مشاهده می‌شود که در رابطه و یا مجرد بودن تاثیر چندانی بر روی مقدار هزینه و خرید مشتریان ندارد و هر دو گروه از مشتریان مقدار تقریباً یکسانی (حدود ۶۰۰ دلار) هزینه داشته‌اند.

بررسی تاثیر تعداد فرزندان بر هزینه‌ها

نمودار میله‌ای برای بررسی تاثیر تعداد فرزندان بر هزینه‌ها به صورت زیر رسم شده است:

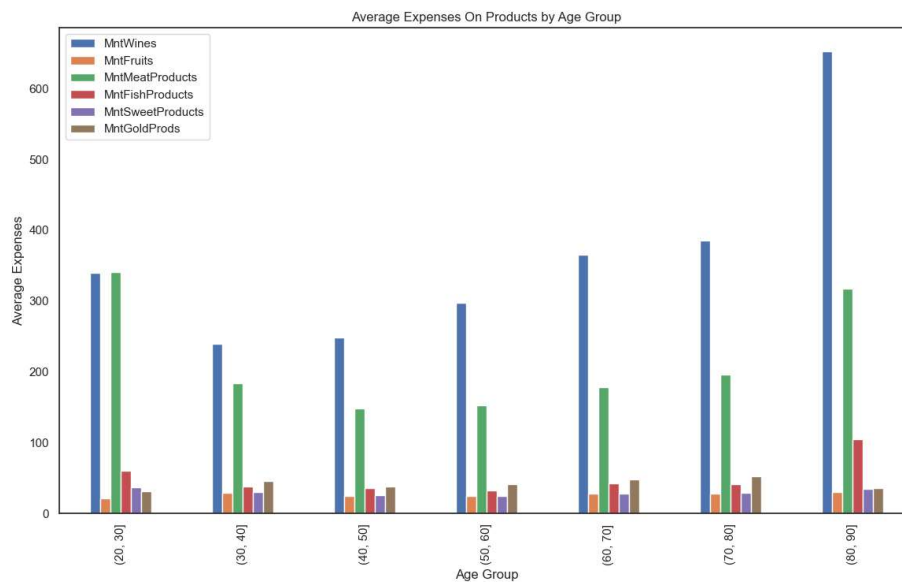


شکل ۱۶- نمودار میله ای تاثیر تعداد فرزندان بر هزینه ها

مشاهده می شود که مشتریانی که هیچ فرزندی ندارند، خرید بسیار بیشتری کرده اند و این موضوع می تواند بیان کننده این باشد که خانواده های دارای فرزند بیشتر باید کنترل و دقت بیشتری بر روی هزینه ها و مقدار خرید خود و مدیریت هزینه های خود برای زندگی باشند.

بررسی میانگین هزینه ها به تفکیک گروه محصولات در گروه های سنی مختلف

نمودار میله ای برای بررسی میانگین هزینه ها به تفکیک گروه محصولات در گروه های سنی مختلف به صورت زیر رسم شده است:

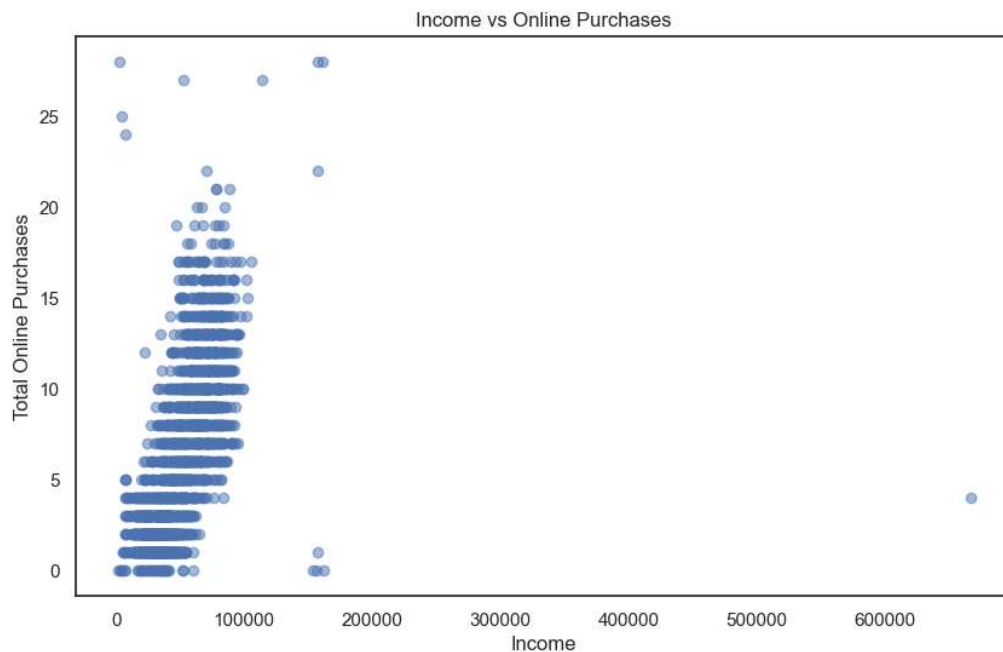


شکل ۱۷- نمودار میله ای میانگین هزینه ها به تفکیک گروه محصولات در گروه های سنی مختلف

در تمام گروه های سنی، هزینه های MntWines به طور قابل توجهی بیشتر از هر دسته محصول دیگری است. این روند نشان دهنده افزایش کلی در هزینه های شراب با افزایش سن است که در گروه سنی ۸۰ تا ۹۰ سال به اوج خود می رسد. MntMeatProducts دارای دومین هزینه در تمام گروه های سنی است. برخلاف شراب، روند آن با افزایش جزئی در گروه های سنی بالاتر، نسبتاً ثابت باقی می ماند. MntSweetProducts و MntGoldProds این دسته ها کمترین هزینه را در تمام گروه های سنی دارند. مخارج آنها نسبتاً ثابت است و روند افزایشی یا کاهشی قوی را نشان نمی دهد. هزینه های شراب در گروه های سنی بالاتر به طور قابل توجهی افزایش می یابد، که نشان دهنده تغییر ترجیح به سمت شراب با افزایش سن افراد است.

بررسی رابطه بین درآمد مشتری و تعداد خریدهای آنلاین

نمودار پراکندگی برای بررسی رابطه بین درآمد مشتری و تعداد خریدهای آنلاین به صورت زیر رسم شده است:

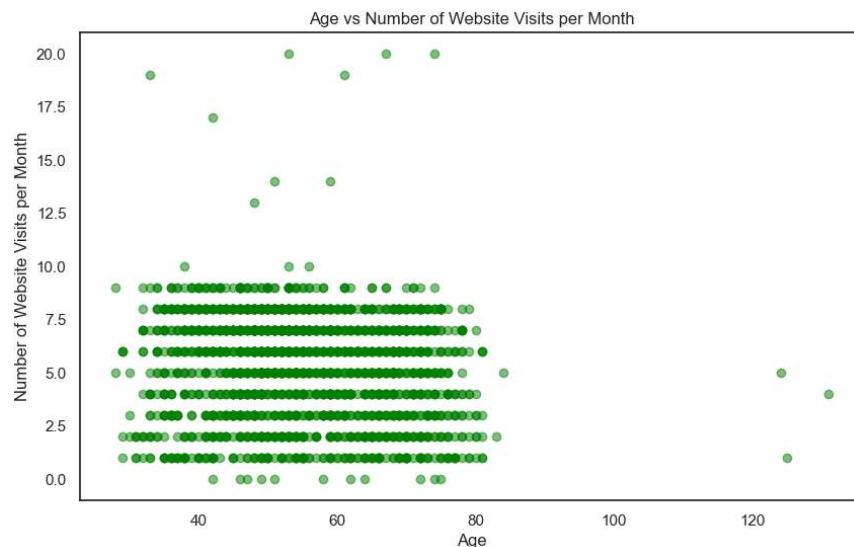


شکل ۱۸- نمودار پراکندگی رابطه بین درآمد مشتری و تعداد خریدهای آنلاین

مشاهده می شود که بین میزان درآمد و خرید های آنلاین رابطه خطی چندانی برقرار نیست اما می توان گفت که رابطه کمی هم بین این دو متغیر برقرار نیست و ارتباطی نسبتاً متوسطی با یکدیگر دارند.

بررسی رابطه سن و تعداد بازدید ماهانه از سایت

نمودار پراکندگی برای بررسی رابطه بین سن مشتری و تعداد بازدید ماهانه از سایت به صورت زیر رسم شده است:

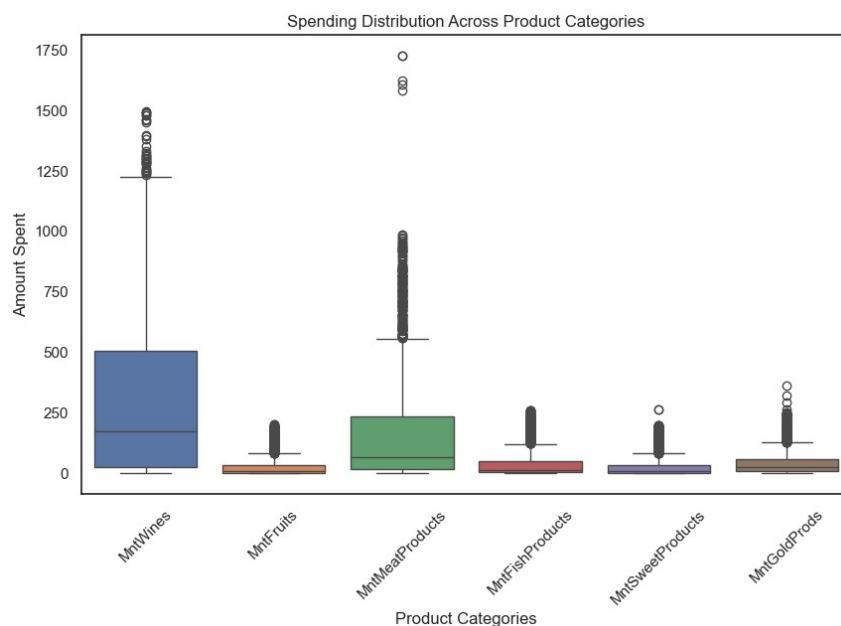


شکل ۱۹- نمودار پراکندگی رابطه بین سن مشتری و تعداد بازدید ماهانه از سایت

مشاهده می‌شود که بین سن و تعداد بازدید ماهانه از سایت رابطه خطی برقرار نیست و این دو متغیر رابطه پایینی با یکدیگر دارند.

بررسی توزیع هزینه به تفکیک گروه محصولات

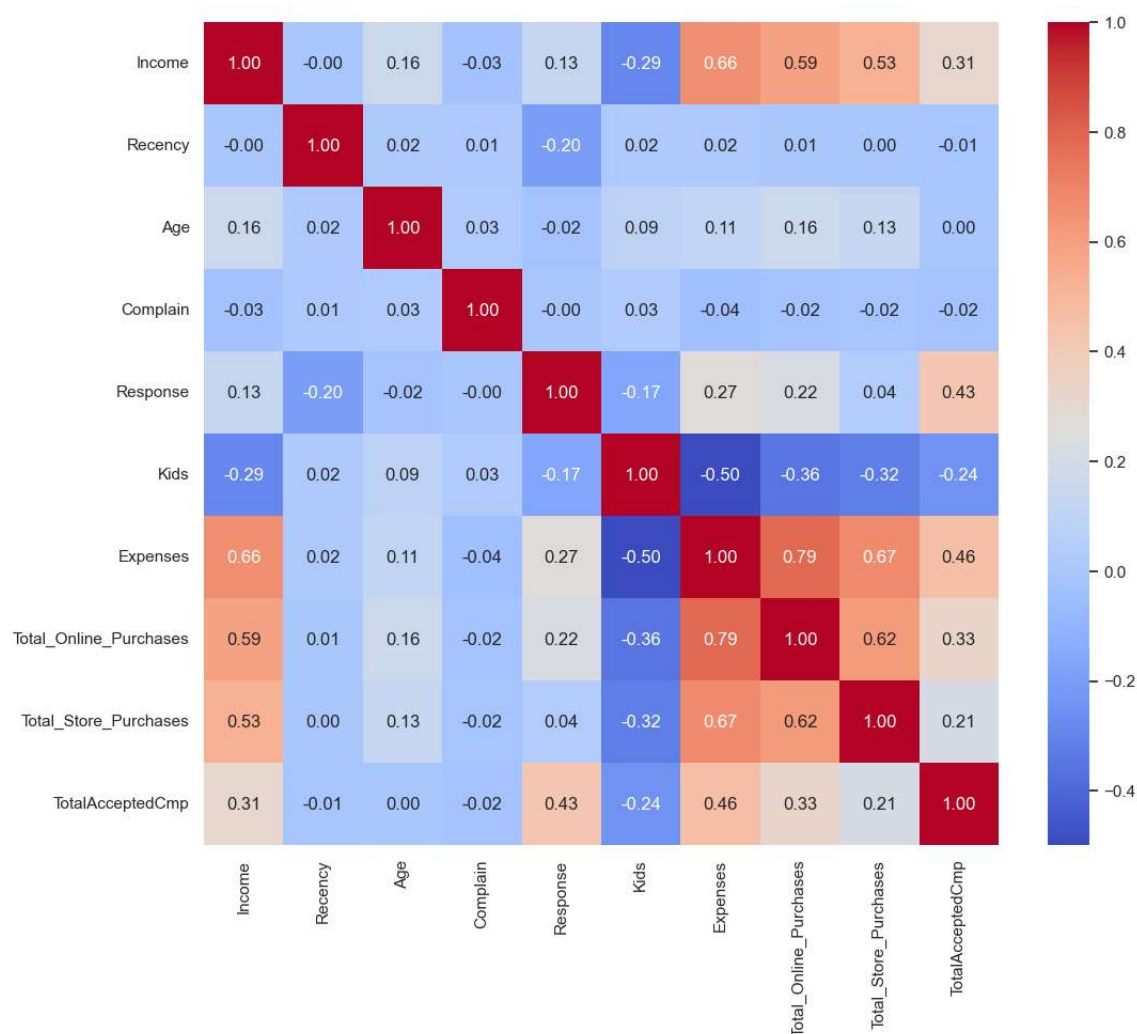
نمودار جعبه ای برای بررسی توزیع هزینه به تفکیک گروه محصولات به صورت زیر رسم شده است:



شکل ۲۰- نمودار جعبه ای توزیع هزینه به تفکیک گروه محصولات

MntWines دارای بالاترین متغیر هزینه است. دسته MntWines دارای بالاترین میانگین هزینه است که نشان دهنده تنوع قابل توجهی در هزینه های مشتریان است. همچنین دارای نقاط دور افتاده بسیاری است که نشان می دهد برخی از مشتریان به طور قابل توجهی بیشتر از میانگین هزینه می کنند. MntMeatProducts نیز تنوع هزینه بالایی را نشان می دهد. دسته MntMeatProducts دومین میانگین هزینه های متوسط و تعداد زیادی موارد دور افتاده را دارد، اگرچه دامنه میان چارکی آن کوچکتر از شراب است. سایر دسته ها هزینه های کمتر و ثابت تری دارند. همه دسته ها برخی از موارد دور افتاده را نشان می دهند، اما آنها در MntWines و MntMeatProducts برجسته تر هستند، جایی که افراد خاص به طور قابل توجهی بیشتر از اکثریت خرج می کنند.

بررسی همبستگی بین متغیر ها



شکل ۲۱- نمودار حرارتی ۱۰ متغیر برای بررسی همبستگی میان آنها

از نمودار حرارتی همبستگی، می‌توان مشاهده کرد که متغیر هزینه (Expenses) با متغیر مجموع خرید های آنلاین دارای بیشترین همبستگی و برابر ۰.۷۹ است که کاملاً علت این موضوع واضح می‌باشد. از دیگر همبستگی های زیاد می‌توان به همبستگی بین دو متغیر Expenses و Total_Store_Purchases اشاره کرد که برابر ۰.۶۷ است. همانطور که قبل تر اشاره کردیم میزان درآمد مشتریان با توان خرید آنها رابطه زیادی را دارد و در نمودار حرارتی مشاهده می‌کنیم که ضریب همبستگی بین آنها برابر ۰.۶۶ است. در کل مشاهده می‌شود که همبستگی آنچنانی بین این متغیر ها وجود ندارند و تنها تعداد کمی از متغیرها دارای ارتباط خوبی هستند.

۳. آزمون های فرض

در این بخش به بررسی و پیاده سازی ۵ آزمون فرض آماری با استفاده از کتابخانه Scipy در زبان برنامه نویسی پایتون پرداخته‌ایم که به چند سوال کلیدی و مهم در این کسب و کار و شناخت رفتار های مشتریان و عوامل موثر بر خرید آنها پاسخ دهیم.

۳/۱. آزمون فرض اول

در این قسمت از تحلیل شخصیت مشتریان، هدف ما بررسی تفاوت میان سطوح تحصیلاتی مشتریان و میزان درآمد آنان است. به منظور انجام این تحلیل، از آزمون Kruskal-Wallis بهره گرفته شد تا تفاوت معنادار در درآمد بین میان گروه‌های تحصیلی ارزیابی شود.

فرض صفر (H_0): میانگین درآمد بین مشتریان با سطوح مختلف تحصیلاتی تفاوتی ندارد.

فرض صفر (H_1): در میان مشتریان با سطوح مختلف تحصیلاتی، درآمدها تفاوت معناداری دارد.

نتایج آزمون: با انجام آزمون Kruskal-Wallis، p-value برابر 8.363×10^{-30} به دست آمد که کمتر از سطح معناداری (معمولاً ۰.۰۵) می‌باشد. این مسئله نشان می‌دهد که فرضیه صف رد می‌شود و می‌توان به این نتیجه رسید که احتمالاً اختلاف معناداری در میزان درآمد میان گروه‌های مختلف تحصیلاتی وجود دارد.

ما می‌توانیم بیان کنیم که تحصیلات به عنوان یک فاکتور مهم در افزایش یا کاهش درآمد نقش دارد.

۳/۲. آزمون فرض دوم

برای بررسی اینکه آیا کمپین بازاریابی تأثیر معناداری بر میزان هزینه کرد مشتریان دارد یا خیر، ابتدا نرمال بودن داده‌ها بررسی شد. آزمون Shapiro-Wilk برای بررسی نرمال بودن متغیر Expenses انجام شد که مقدار p -value برابر 2.78×10^{-40} را نشان داد. با توجه به مقدار p -value کمتر از ۰.۰۵، فرض نرمال بودن داده‌ها رد شد.

از آنجا که داده‌ها توزیع نرمال نداشتند، از آزمون Mann-Whitney U برای مقایسه هزینه بین دو گروه مشتریانی که به کمپین بازاریابی پاسخ مثبت داده‌اند ($\text{Response} = ۱$) و مشتریانی که به آن پاسخ نداده‌اند ($\text{Response} = ۰$) استفاده شد.

نتیجه آزمون نشان داد که مقدار p -value برابر 1.44×10^{-29} بوده که کمتر از ۰.۰۵ است. بنابراین، کمپین بازاریابی تأثیر معناداری بر رفتار هزینه‌ای مشتریان دارد و مشتریانی که به کمپین پاسخ داده‌اند، الگوی هزینه‌ای متفاوتی نسبت به سایر مشتریان دارند.

۳/۳. آزمون فرض سوم

برای بررسی اینکه آیا مشتریانی که فرزند دارند نسبت به مشتریانی که فرزندی ندارند الگوی هزینه‌ای متفاوتی دارند یا خیر، ابتدا یک متغیر جدید به نام Has_Children ایجاد شد. این متغیر مشخص می‌کند که آیا مشتری حداقل یک فرزند در خانه دارد یا خیر (با جمع کردن تعداد کودکان و نوجوانان در خانه). با توجه به اینکه داده‌ها نرمال نبودند، از آزمون Mann-Whitney U برای مقایسه هزینه بین دو گروه دارای فرزند و بدون فرزند استفاده شد.

نتیجه آزمون نشان داد که مقدار p -value برابر 7.30×10^{-108} بوده که بسیار کمتر از ۰.۰۵ است. بنابراین، تفاوت معناداری در میزان هزینه کرد مشتریان دارای فرزند و بدون فرزند وجود دارد و این دو گروه الگوهای هزینه‌ای متفاوتی دارند.

۳/۴. آزمون فرض چهارم

برای بررسی اینکه آیا میزان هزینه‌ای که مشتریان برای دسته‌بندی‌های مختلف محصولات صرف می‌کنند متفاوت است یا خیر، از آزمون Friedman استفاده شد. این آزمون برای مقایسه گروه‌های مرتبط در داده‌های وابسته به کار می‌رود. دسته‌بندی‌های بررسی‌شده شامل موارد زیر بودند:

- شراب (MntWines)
- میوه‌ها (MntFruits)
- محصولات گوشتی (MntMeatProducts)
- محصولات دریایی (MntFishProducts)
- شیرینی‌جات (MntSweetProducts)
- محصولات طلایی (MntGoldProds)

نتایج آزمون نشان داد که مقدار p -value برابر صفر بوده که کمتر از ۰.۰۵ است. بنابراین، تفاوت معناداری در میزان هزینه‌کرد مشتریان بین دسته‌بندی‌های مختلف محصولات وجود دارد. این یافته نشان می‌دهد که مشتریان برای برخی از محصولات بیشتر هزینه می‌کنند و این توزیع هزینه بین دسته‌های مختلف یکنواخت نیست.

۳/۵. آزمون فرض پنجم

برای بررسی اینکه آیا بین سطح تحصیلات مشتریان و پذیرش کمپین‌های تبلیغاتی ارتباط معناداری وجود دارد یا خیر، از آزمون χ^2 (Chi-Square Test) استفاده شد. این آزمون برای بررسی وابستگی بین دو متغیر کیفی به کار می‌رود. ابتدا، یک Contingency Table بین دو متغیر "تحصیلات" و "پاسخ به کمپین تبلیغاتی" ساخته شد. سپس آزمون χ^2 روی این جدول اجرا شد و مقدار p -value = ۰.۰۰۰۱۲ به دست آمد که کمتر از ۰.۰۵ است.

از آنجایی که p -value کمتر از ۰.۰۵ است، می‌توان نتیجه گرفت که سطح تحصیلات مشتریان به‌طور معناداری با پذیرش کمپین‌های تبلیغاتی مرتبط است. این یافته نشان می‌دهد که میزان تأثیرگذاری کمپین‌های تبلیغاتی ممکن است بسته به سطح تحصیلات مشتریان متفاوت باشد.