**Deliverable 1:**

Table

Description automatically generated

از خروجی بالا می بینیم که:

1. طول وسیله نقلیه و فاصله وسیله نقلیه از زمین از نظر آماری به احتمال زیاد مقادیر غیر تصادفی واریانس مدل را فراهم می کند. یعنی طول وسیله نقلیه و فاصله وسیله نقلیه از زمین تأثیر قابل توجهی بر مایل در هر گالن در نمونه اولیه MechaCar دارد. برعکس، وزن وسیله نقلیه، زاویه اسپویلر و تمام چرخ محرک (AWD) دارای مقادیر p-مقدارهایی هستند که میزان واریانس تصادفی با مجموعه داده را نشان می دهد.
2. p-Value برای این مدل، p-Value: 5.35e-11، بسیار کوچکتر از سطح معناداری فرضی 0.05٪ است. این نشان می دهد که شواهد کافی برای رد فرضیه صفر ما وجود دارد، که بیشتر نشان می دهد که شیب این مدل خطی صفر نیست.
3. این مدل خطی دارای مقدار r-squared 0.7149 است که به این معنی است که تقریباً 71٪ از کل پیش بینی های mpg توسط این مدل تعیین می شود. به طور نسبی، مدل رگرسیون چندگانه او به طور موثر mpg نمونه های اولیه MechaCar را پیش بینی می کند.

اگر متغیرهای مستقل کمتر تاثیرگذار (وزن خودرو، زاویه اسپویلر و تمام چرخ محرک) را حذف کنیم، قابل پیش بینی کاهش می یابد، اما نه به شدت: مقدار r-squared از 0.7149 به 0.674 کاهش می یابد.

Text

Description automatically generated

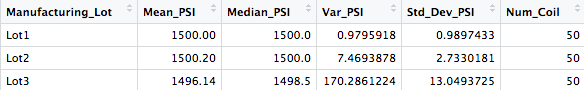
# Deliverable 2:

مجموعه داده کویل تعلیق ارائه شده برای MechaCar شامل نتایج آزمایش ظرفیت وزنی کویل های تعلیق متعدد از لات های تولیدی متعدد برای تعیین سازگاری است.

ابتدا نگاهی به تمام بخش های تولیدی:



با درک این موضوع که مشخصات طراحی برای سیم‌پیچ‌های تعلیق MechaCar الزام می‌کند که واریانس سیم‌پیچ‌های تعلیق نباید از 100 پوند بر اینچ مربع (PSI) تجاوز کند.



هنگامی که به کل جمعیت لات تولید نگاه می کنیم، واریانس کویل ها 62.29 PSI است که به خوبی در حد نیاز واریانس 100 PSI است.

به طور مشابه، اما به طور قابل توجهی سازگارتر، Lot 1 و Lot 2 به خوبی در شرایط واریانس 100 PSI هستند. با واریانس 0.98 و 7.47 به ترتیب. با این حال، لات 3 است که واریانس بسیار بیشتری را در عملکرد و ثبات نشان می دهد، با واریانس 170.29. این لات 3 است که به طور نامتناسبی باعث ایجاد واریانس در سطح لات کامل می شود.

این باکس نمودار بسیار ساده تفاوت های بین لات ها را نشان می دهد:

Chart, box and whisker chart

Description automatically generated

# Deliverable 3:

گام بعدی انجام یک آزمون t بر روی داده‌های کویل تعلیق است تا مشخص شود که آیا تفاوت آماری بین میانگین مجموعه داده‌های نمونه ارائه شده و مجموعه داده‌های جمعیتی فرضی وجود دارد یا خیر. با استفاده از میانگین جمعیت فرضی 1500، موارد زیر را می یابیم:

خلاصه ای از نتایج آزمون t در تمام لات های تولیدی وجود دارد

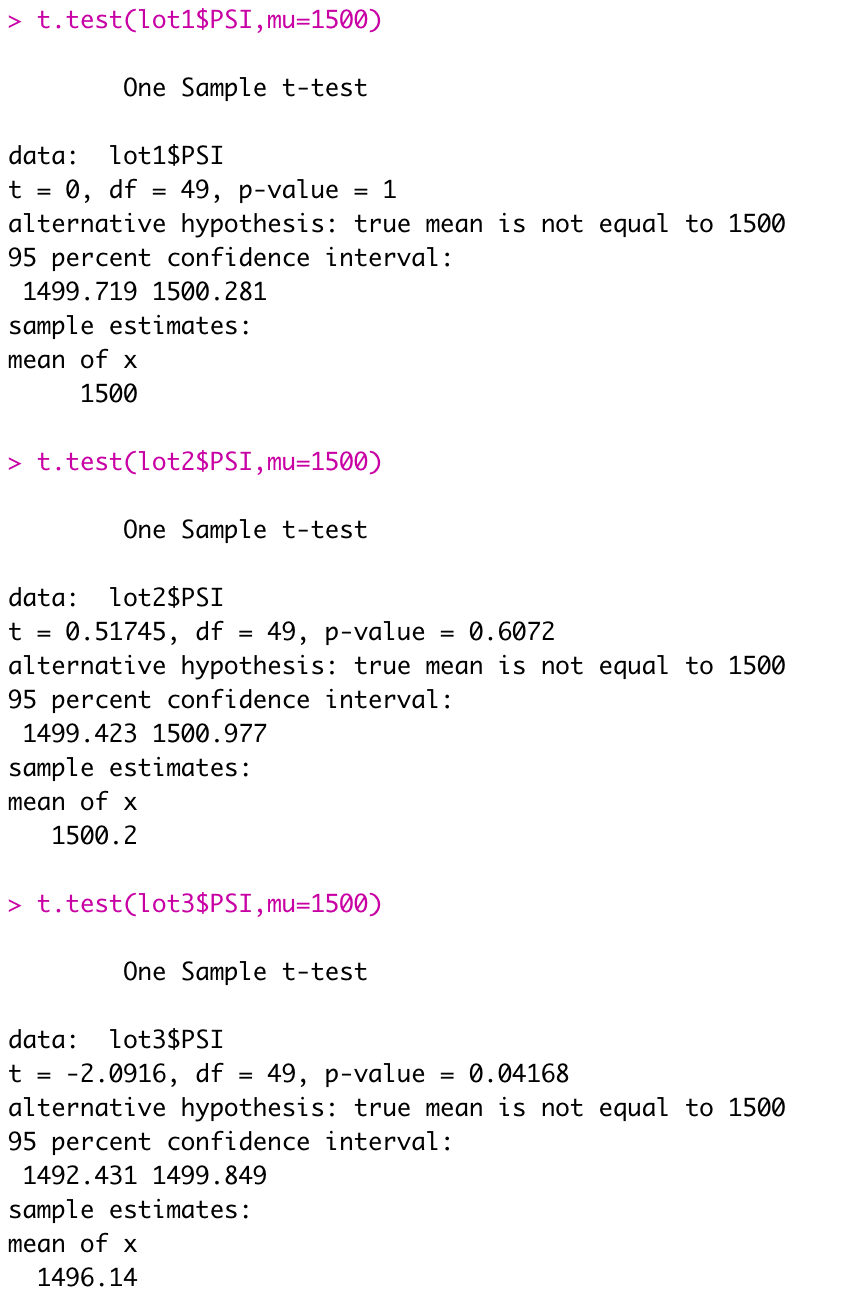
Text, letter

Description automatically generated

در اینجا می توانیم میانگین واقعی نمونه را 1498.78 ببینیم که در آمار خلاصه بالا نیز مشاهده کردیم. با p-Value 0.06، که بالاتر از سطح معنی‌داری رایج 0.05 است، شواهد کافی برای رد فرضیه صفر وجود ندارد. یعنی میانگین هر سه قطعه تولیدی مذکور از نظر آماری مشابه میانگین جمعیت فرضی 1500 است.

نگاه بعدی به هر بخش جداگانه:

1. نمونه لات 1 در واقع دارای میانگین نمونه واقعی 1500 است، همانطور که در آمار خلاصه بالا دیدیم. با p-Value 1، به وضوح نمی توانیم فرضیه صفر را رد کنیم (یعنی بپذیریم) که هیچ تفاوت آماری بین میانگین نمونه مشاهده شده و میانگین جامعه فرضی (1500) وجود ندارد.
2. لات 2 اساساً همان نتیجه را با میانگین نمونه 1500.02 دارد، مقدار p-Value 0.61. فرضیه صفر را نمی توان رد کرد و میانگین نمونه و میانگین جامعه 1500 از نظر آماری مشابه هستند.
3. با این حال، لات 3، جای تعجب نیست که سناریوی متفاوتی باشد. در اینجا میانگین نمونه 1496.14 و p-Value 0.04 است که کمتر از سطح معنی داری رایج 0.05 است. همه حاکی از رد فرضیه صفر است که میانگین این نمونه و میانگین جامعه فرضی از نظر آماری متفاوت نیستند.



این اطلاعات چگونه به شما کمک می کند؟ واضح است که چیزی در چرخه تولید Lot 3 به هم ریخته است. فرآیند باید از نظر خرابی سیستم بررسی شود و پیچ های تعلیق این لات باید بازرسی شود تا مواردی که کیفیت های کیفیت را حذف نمی کنند.

# Deliverable 4:

معیارهای

جمع‌آوری داده‌ها برای مدل‌های قابل مقایسه در همه تولیدکنندگان بزرگ در ۳ سال گذشته برای معیارهای زیر:

رتبه بندی ویژگی های ایمنی: متغیر مستقل

قیمت فعلی (فروش): متغیر وابسته

بسته Drive: متغیر مستقل

موتور (برقی، هیبریدی، بنزینی / معمولی): متغیر مستقل

ارزش فروش مجدد: متغیر مستقل

میانگین سالانه هزینه مالکیت (نگهداری): متغیر مستقل

MPG (بازده بنزین): متغیر مستقل

Hypothesis: Null and Alternative.

نپس از تعیین عوامل کلیدی برای ژانر MechaCar:

. Null Hypothesis (Ho) : MechaCar بر اساس عملکرد آن از عوامل کلیدی برای ژانر خود، قیمت گذاری درستی دارد.

: Alternative Hypothesis (Ha) MechaCar بر اساس عملکرد عوامل کلیدی ژانر خود به درستی قیمت گذاری نمی شود.

آزمون های آماری

یک رگرسیون خطی چندگانه برای تعیین عواملی که بیشترین همبستگی/قابلیت پیش بینی را با لیست قیمت فروش دارند (متغیر وابسته) استفاده می شود. کدام ترکیب بیشترین تأثیر را روی قیمت دارد (ممکن است همه آنها باشند!)