**Assignment on R module**

**Submitted by;**

Saddam Hossain

DS-2505 batch

**Submitted to**

Naimul Hassan

Chief technology officer and data Scientist

**Works on “mtcars” data set default in R**

**Qustions**

* **The distribution and relationships between variables relevant to MPG.**
* **The strength of these relationships with Pearson’s correlation.**
* **Which variables have the strongest positive and negative relationships with MPG.**

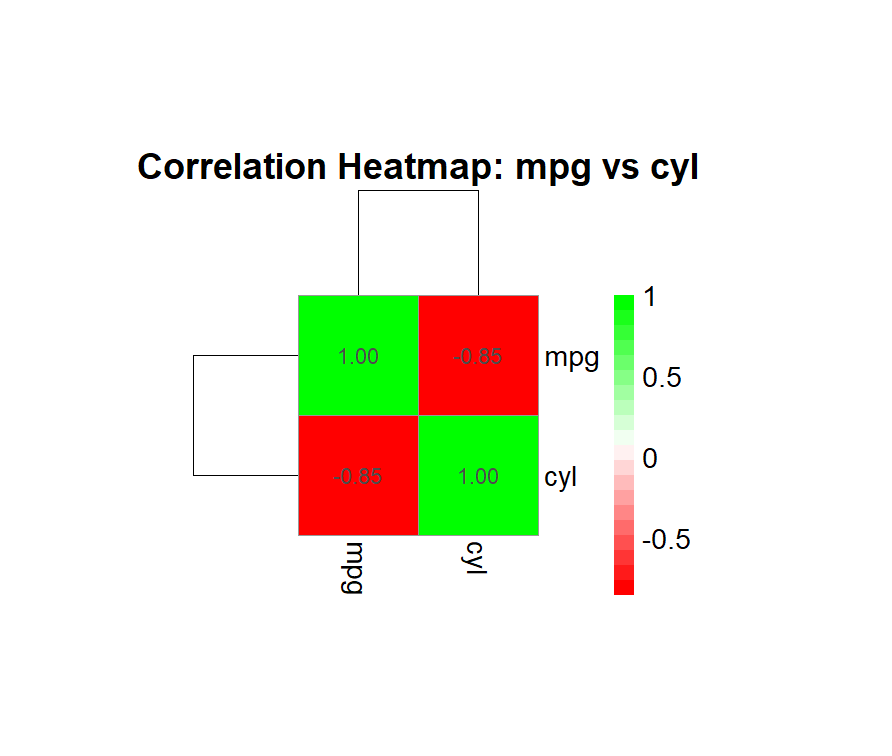
**Solutions**

**Distribution and relationship between variables mpg and cyl**

Here, mpg stands for (miles per gallon = fuel efficiency) and cyl = (number of cylinders).

**Key Results**

* **Correlation coefficient (cor)**: **-0.852**  
  → This is a **strong negative correlation** between MPG and number of cylinders.  
  Meaning: as the number of cylinders increases, fuel efficiency decreases sharply.
* **t-statistic (t = -8.9197) & df = 30**  
  → The test statistic is large in magnitude, showing the relationship is far from random.
* **p-value = 6.113e-10** (≈ 0.0000000006)  
  → Extremely small, much less than any standard significance level (0.05, 0.01, or 0.001).  
  We **reject the null hypothesis** (which assumes no correlation).  
  So, the negative relationship is statistically significant.
* **95% Confidence Interval: [-0.926, -0.716]**  
  → We are 95% confident that the true population correlation lies between -0.93 and -0.72.  
  This confirms the correlation is not just negative, but **strongly negative**.

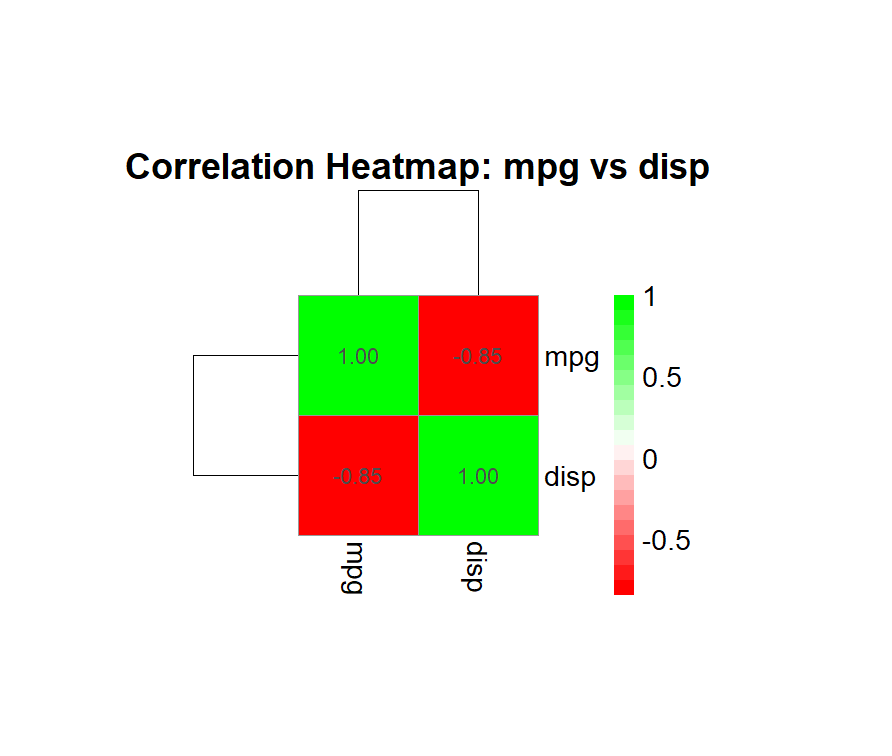


**Distribution and relationship between variables mpg and disp**

Here, disp stands for Displacement.

**Key Results**

* **Correlation coefficient (cor): -0.848**  
  → This is a **strong negative correlation** between fuel efficiency (MPG) and engine displacement.  
  **Meaning:** as engine displacement increases, fuel efficiency decreases sharply.
* **t-statistic (t = -8.7472) & df = 30**  
  → The test statistic is large in magnitude, showing the relationship is far from random.
* **p-value = 9.38e-10 (≈ 0.0000000009)**  
  → Extremely small, much less than any standard significance level (0.05, 0.01, or 0.001).  
   We reject the null hypothesis (which assumes no correlation).  
   So, the negative relationship is **statistically significant**.
* **95% Confidence Interval: [-0.923, -0.708]**  
  → We are 95% confident that the true population correlation lies between **-0.92 and -0.71**.  
  This confirms the correlation is not just negative, but **strongly negative**.



**Key Results: Correlations with MPG**

1. **mpg vs hp**

* **Correlation coefficient (cor): -0.776**  
  → **Strong negative correlation**: as horsepower increases, fuel efficiency decreases.
* **t-statistic (t = -6.7424) & df = 30**  
  → Large magnitude, indicating relationship is far from random.
* **p-value = 1.79e-07**  
  → **Extremely small**; statistically significant.
* **95% CI: [-0.885, -0.586]**  
  → Strong negative correlation is confirmed.

1. **mpg vs drat**

* **Correlation coefficient: 0.681**  
  → **Strong positive correlation:** higher rear axle ratio tends to increase MPG.
* **t = 5.096, df = 30, p-value = 1.78e-05**  
  → Statistically significant positive relationship.
* **95% CI: [0.436, 0.832]**

1. **mpg vs wt**

* **Correlation coefficient: -0.868**  
  → Very strong negative correlation: heavier cars have lower fuel efficiency.
* **t = -9.559, df = 30, p-value = 1.29e-10**  
  → Highly significant.
* **95% CI: [-0.934, -0.744]**

1. **mpg vs qsec**

* **Correlation coefficient: 0.419**  
  → Moderate positive correlation: higher 1/4 mile time slightly associated with better MPG.
* **t = 2.5252, df = 30, p-value = 0.017**  
  → Statistically significant at 5% level.
* **95% CI: [0.082, 0.67]**

1. **mpg vs vs**

* **Correlation coefficient: 0.664**  
  → Strong positive correlation: cars with V/S engine type = 1 tend to have higher MPG.
* **t = 4.8644, df = 30, p-value = 3.42e-05**  
  → Statistically significant.
* **95% CI: [0.41, 0.822]**

1. **mpg vs am**

* **Correlation coefficient: 0.6**  
  → Strong positive correlation: cars with manual transmission tend to have better fuel efficiency.
* **t = 4.1061, df = 30, p-value = 0.000285**  
  → Statistically significant.
* **95% CI: [0.318, 0.784]**

1. **mpg vs gear**

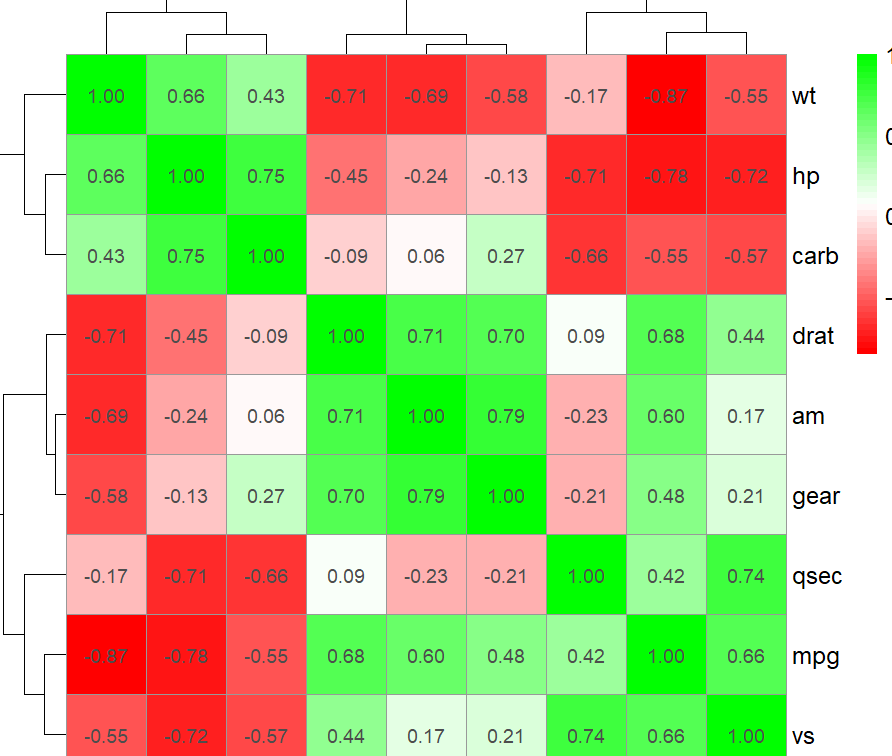
* **Correlation coefficient: 0.48**  
  → Moderate positive correlation: cars with more gears slightly tend to have higher MPG.
* **t = 2.9992, df = 30, p-value = 0.0054**  
  → Statistically significant.
* **95% CI: [0.158, 0.71]**

1. **mpg vs carb**

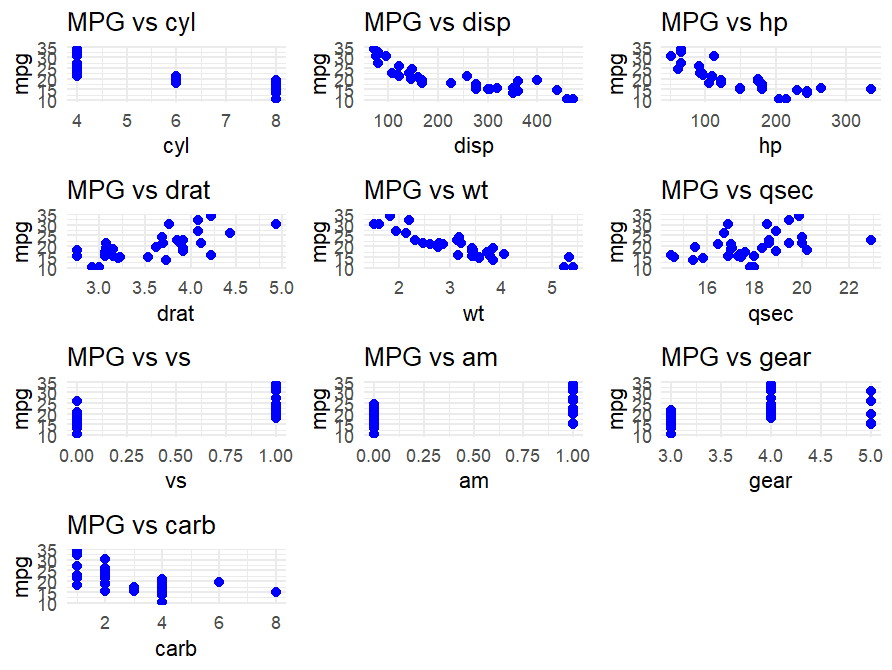
* **Correlation coefficient: -0.551**  
  → Moderate negative correlation: cars with more carburetors tend to have lower MPG.
* **t = -3.6157, df = 30, p-value = 0.00108**  
  → Statistically significant.
* **95% CI: [-0.755, -0.25]**

**🔹 Summary Insights**

* **Strong negative correlations:** wt (-0.868), hp (-0.776), carb (-0.551) → bigger, heavier, and more powerful cars are less fuel-efficient.
* **Strong positive correlations:** drat (0.681), vs (0.664), am (0.6) → cars with higher rear axle ratio, V/S engine, and manual transmission are more fuel-efficient.
* **Moderate correlations:** qsec (0.419), gear (0.48) → smaller but still meaningful associations.
* All correlations are **statistically significant** at p < 0.05.



**Scatter plot between the variables**



**MPG (miles per gallon) vs other variables**, here are some key insights:

**1. MPG vs Cylinders (cyl)**

* Negative relationship: Cars with more cylinders tend to have lower MPG.
* Most 4-cylinder cars have higher MPG (~25–35), while 8-cylinder cars are mostly low (~10–15 MPG).

**2. MPG vs Displacement (disp)**

* Negative correlation: Higher engine displacement generally leads to lower MPG.
* Cars with small displacement engines (~70–150) have high MPG (~25–35), while large engines (~300–400) have low MPG (~10–20).

**3. MPG vs Horsepower (hp)**

* Negative relationship: More horsepower usually means lower fuel efficiency.
* High horsepower (>200) cars are typically low on MPG (<20), whereas lower horsepower (<100) cars achieve higher MPG.

**4. MPG vs Rear Axle Ratio (drat)**

* Positive trend: Higher drat seems to correspond to higher MPG.
* Most cars with drat > 4 have higher MPG.

**5. MPG vs Weight (wt)**

* Strong negative correlation: Heavier cars have lower MPG.
* Light cars (~1.5–2.5) achieve high MPG, whereas heavy cars (>4) fall below 15 MPG.

**6. MPG vs 1/4 Mile Time (qsec)**

* Weak relationship: MPG does not strongly depend on quarter-mile time.
* MPG varies widely across qsec values (~14–22 sec).

**7. MPG vs Engine Shape (vs)**

* Categorical variable (0 = V-shaped, 1 = Straight)
* Straight engines (vs = 1) tend to have higher MPG.

**8. MPG vs Transmission (am)**

* Categorical variable (0 = automatic, 1 = manual)
* Manual transmission cars (am = 1) tend to have higher MPG.

**9. MPG vs Number of Gears (gear)**

* Slight trend: Cars with more gears (4–5) may have slightly higher MPG.
* 3-gear cars are mostly low on MPG.

**10. MPG vs Number of Carburetors (carb)**

* Negative trend: Cars with more carburetors tend to have lower MPG.
* 1–2 carb cars have high MPG; 4–8 carb cars generally have lower MPG.

**Summary Insights**

* **Strongest negative correlations with MPG:** cyl, disp, hp, wt, carb.
* **Positive contributors to MPG:** drat, vs = 1 (straight engines), am = 1 (manual), more gears (slight effect).
* **Weak relationship:** qsec (1/4 mile time).