R Notebook

Code

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Cmd+Shift+Enter*.

Hide

```
library(dplyr)
mechaCar_mpg_df <- read.csv('MechaCar_mpg.csv' , check.names=F, stringsAsFactors = F)
head(mechaCar_mpg_df)
```

	vehicle_length <dbl>	vehicle_weight <dbl>	spoiler_angle <dbl>	ground_clearance <dbl>	A... <int>	mpg <dbl>
1	14.69710	6407.946	48.78998	14.64098	1	49.04918
2	12.53421	5182.081	90.00000	14.36668	1	36.76606
3	20.00000	8337.981	78.63232	12.25371	0	80.00000
4	13.42849	9419.671	55.93903	12.98936	1	18.94149
5	15.44998	3772.667	26.12816	15.10396	1	63.82457
6	14.45357	7286.595	30.58568	13.10695	0	48.54268

6 rows

Hide

```
lm(mpg ~ vehicle_length + vehicle_weight + spoiler_angle + ground_clearance + AWD,data=mechaCar_mpg_df) #generate multiple linear regression model
```

```

Call:
lm(formula = mpg ~ vehicle_length + vehicle_weight + spoiler_angle +
    ground_clearance + AWD, data = mechaCar_mpg_df)

Coefficients:
     (Intercept)    vehicle_length    vehicle_weight
      -1.040e+02         6.267e+00         1.245e-03
   spoiler_angle  ground_clearance               AWD
       6.877e-02         3.546e+00        -3.411e+00
```

Hide

```
summary(lm(mpg ~ vehicle_length + vehicle_weight + spoiler_angle + ground_clearance + AWD,data=mechaCar_mpg_df))
#generate summary statistics
```

```

Call:
lm(formula = mpg ~ vehicle_length + vehicle_weight + spoiler_angle +
    ground_clearance + AWD, data = mechaCar_mpg_df)

Residuals:
    Min      1Q  Median      3Q     Max
-19.4701  -4.4994  -0.0692   5.4433  18.5849

Coefficients:
                  Estimate Std. Error t value
(Intercept)     -1.040e+02  1.585e+01  -6.559
vehicle_length   6.267e+00  6.553e-01   9.563
vehicle_weight   1.245e-03  6.890e-04   1.807
spoiler_angle    6.877e-02  6.653e-02   1.034
ground_clearance 3.546e+00  5.412e-01   6.551
AWD             -3.411e+00  2.535e+00  -1.346
                Pr(>|t|)
(Intercept)     5.08e-08 ***
vehicle_length  2.60e-12 ***
vehicle_weight     0.0776 .
spoiler_angle      0.3069
ground_clearance 5.21e-08 ***
AWD                0.1852
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.774 on 44 degrees of freedom
Multiple R-squared:  0.7149,    Adjusted R-squared:  0.6825
F-statistic: 22.07 on 5 and 44 DF,  p-value: 5.35e-11
```

Hide

```
#Summary Statistics on Suspension Coils
suspension_coil_table <- read.csv('Suspension_Coil.csv' , check.names=F, stringsAsFactors = F)
head(suspension_coil_table)
```

	VehicleID <chr>	Manufacturing_Lot <chr>	PSI <int>
1	V40858	Lot1	1499
2	V40607	Lot1	1500
3	V31443	Lot1	1500
4	V6004	Lot1	1500
5	V7000	Lot1	1501
6	V17344	Lot1	1501

6 rows

Hide

```
total_summary_df <- suspension_coil_table %>% summarize(Mean=mean(PSI), Median=median(PSI), Variance=var(PSI), SD
=sd(PSI), .groups = 'keep') #create summary table
#head(total_summary_df)
total_summary_df
```

Mean <dbl>	Median <dbl>	Variance <dbl>	SD <dbl>
1498.78	1500	62.29356	7.892627

1 row

Hide

```
lot_summary <- suspension_coil_table %>% group_by(Manufacturing_Lot) %>% summarize(Mean=mean(PSI), Median=median(
PSI), Variance=var(PSI), SD=sd(PSI), .groups = 'keep') #create summary table
lot_summary
```

Manufacturing_Lot <chr>	Mean <dbl>	Median <dbl>	Variance <dbl>	SD <dbl>
Lot1	1500.00	1500.0	0.9795918	0.9897433
Lot2	1500.20	1500.0	7.4693878	2.7330181
Lot3	1496.14	1498.5	170.2861224	13.0493725

3 rows

Hide

```
NA
```

Hide

```
#T-Test on Suspension Coils
# compare sample versus population means
t.test(suspension_coil$PSI,mu=1500)
```

```

    One Sample t-test

data:  suspension_coil$PSI
t = -1.8931, df = 149, p-value = 0.06028
alternative hypothesis: true mean is not equal to 1500
95 percent confidence interval:
 1497.507 1500.053
sample estimates:
mean of x
  1498.78
```

Hide

```
#2. Use t.test() function 3 more times with subset() to determine if PSI for each manufacturing lot is statistica
lly different from the pop. mean of 1,500 PSI
lot1 <- subset(suspension_coil, Manufacturing_Lot=="Lot1")
lot2 <- subset(suspension_coil, Manufacturing_Lot=="Lot2")
lot3 <- subset(suspension_coil, Manufacturing_Lot=="Lot3")

t.test(lot1$PSI,mu=1500)
```

```

    One Sample t-test

data:  lot1$PSI
t = 0, df = 49, p-value = 1
alternative hypothesis: true mean is not equal to 1500
95 percent confidence interval:
 1499.719 1500.281
sample estimates:
mean of x
     1500
```

Hide

```
t.test(lot2$PSI,mu=1500)
```

```

    One Sample t-test

data:  lot2$PSI
t = 0.51745, df = 49, p-value = 0.6072
alternative hypothesis: true mean is not equal to 1500
95 percent confidence interval:
 1499.423 1500.977
sample estimates:
mean of x
   1500.2
```

Hide

```
t.test(lot3$PSI,mu=1500)
```

```

    One Sample t-test

data:  lot3$PSI
t = -2.0916, df = 49, p-value = 0.04168
alternative hypothesis: true mean is not equal to 1500
95 percent confidence interval:
 1492.431 1499.849
sample estimates:
mean of x
  1496.14
```

Hide

```
# log10 was not used as the data was not skewed and did not need any smoothing
```

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Cmd+Option+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Cmd+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.