

## Week 2 - AYU - Pod

---

### Questions

In this AYU, we will examine the hospital cost of patients in Wisconsin in 2003 using patients' information such as their age, gender, and their length of stay at the hospital.

1. Regress the total charge (TOTCHG) on age and length of stay. What is the R-squared of the regression? With significant level of 5%, is there any variable not significant based on coefficient p-values of the model?

```
# Include the codes here
```

Answer:

2. Use the model to predict the total charge of a 13-year-old that stays a week at a hospital.

```
# Include the codes here
```

Answer:

3. Adding variable GENDER to the regression. Write the equations of this new linear model. Use the model to predict the total charge of a 13-year-old female that stays a week at a hospital.

```
# Include the codes here
```

Answer:

4. Are both AGE and APRDRG not significant and should not be added to the model? Use the F-test to address the question.

```
# Include the codes here
```

Answer:

5. Could you improve the model in 3 using the methods discussed for the term life dataset?

```
# Include the codes here
```

Answer:

6. Make a correlation plot. Calculate the vif. Should one pursue multilinearity analysis on the model?

```
# Include the codes here
```

Answer:

7. In this question, we study the the seatpos dataset to study the car seat position of the driver at their comfort. This dataset is collected by the University of Michigan collected data on 38 drivers.

<https://search.r-project.org/CRAN/refmans/faraway/html/seatpos.html>

We would like to regress **hipcenter** on all other variables. Check the VIF of this regression and handle the multilinearity issue if it occurs.

Answer: