

# Homework 2

---

**Due: Wednesday February 15 at 11:59pm**

See general homework tips and submit your files via the course website.

For exercises 1 and 2 use the **diy** data set defined in **HW2Data.sas** in the Homework 2 folder on the course website, which is a modification of a data set we will use in Chapter 8 (you can read more of the background of the data there).

The data set in **HW2Data.sas** contains the following variables:

- **work** : type of work (**Skilled**, **Unskilled** or **Office**)
- **agegrp** : age group with cutoff at 30 (**youngest** or **older**)
- **response**: **yes** or **no**
- **n**: count for particular response

The **yes** and **no** responses were to the question of whether the individual did home improvements themselves in the previous year that they would have previously hired someone to do.

Exercise 3 uses the iris data set from the SAS help (e.g. **data=sashelp.iris**).

## Exercise 1:

- a) Construct a contingency table for **agegrp** and **response** and comment on any apparent associations between age group and decision to hire someone to do home improvements that the individual would have previously done themselves.
- b) Perform and comment on appropriate tests of association, and interpret the results.
- c) Test (using risk differences) if individuals under 30 had a significantly higher probability than individuals over 30 of doing home improvements they would have previously hired someone to do.

## Exercise 2:

Now perform an analysis based on work type.

- a) Construct a contingency table for **work** and **response** and comment on any apparent associations between work type and decision to hire someone to do home improvements that the individual would have previously done themselves.
- b) Perform and comment on appropriate tests of association, and interpret the results.

## Exercise 3:

Consider a one-way ANOVA model to identify differences between group petal length means. We saw in Homework 1 that the normality assumption is reasonable, so you can proceed without testing normality.

- a) Perform a one-way ANOVA for **petallength** with **species** as the categorical predictor, and test any assumptions of the model that should be tested (aside from normality, which you do not need to test).
- b) Comment on the significance of the model, and the variation described by the model. What do these results indicate about the practical usefulness of the model?
- c) Comment on any statistically significantly different petal length means as determined by the best test for comparing all pairwise differences, and explain what that tells us about differences between species.