

Instructions

Upload your own **solutions** to the corresponding Canvas Assignment on or before 17/12/2021.

Grading basis:
Complete/incomplete

Applying SPSS

Use SPSS to create the tables as mentioned below. Submit via Canvas on the Assignment that corresponds to use of SPSS (do not submit together with Homework #4).

Data. The file “FollowUpSurveySPSS2.sav” contains the full data set on which part **c** of Homework #4 is based. For all 100 employees that have been asked about their vaccination status in the follow-up survey, the variable **Status** encodes the vaccination status as follows: 0 for *no vaccination*, 1 for *one out of two shots*, 2 for *fully vaccinated without booster* and 3 for *fully vaccinated with booster*. In addition to this, the data set also contains the variable **gender** that takes the values *M* and *F* for male and female, respectively, as well as the variable **weight**, which gives the weight gain of the employees during the Corona pandemic in kg (negative numbers correspond to a weight loss). The variable **Survey** indicates whether the employee participated in the initial survey of part **b** (0 for “no”, 1 for “yes”).

Instructions.

1. Open the file “FollowUpSurveySPSS2.sav” on Canvas with SPSS.
2. Our first aim is to check the calculations of part **c** of the assignment and to see how to use cross tables and Chi-square tests in SPSS. To do this, go to

Analyze → *Descriptive Statistics* → *Crosstabs*.

Choose the variable **Survey** in the rows and the variable **Status** in the columns. In the *Cells*-menu select the **expected** and **observed** counts.

► **Report the Cross table with the required counts**

and

- ▷ explain why we cannot apply the χ^2 -test in this case.

3. We will now reduce the number of categories of the variable **Status**. To do this, go to

Transform → *Recode into different variables*.

Choose “Vaccination Status” and “StatusJane” (use your own first name, not Jane!) as new name/label. Then choose Old and New Values and set 0 to 0 (not vaccinated) and set 1,2 and 3 to 1 (vaccinated at least once). After creating **StatusJane**, give the new values 0 and 1 the correct labels via Variable view.

Now in Crosstabs, replace **Status** by **StatusJane**, choose Chi-square in the **Statistics menu** and in the **Cells menu** select the expected and observed counts.

► **Report the Cross table with the required counts and the table with Chi-square tests**

and

- ▷ explain what conclusion you can draw from the result (at a 5% level of significance).

4. We will now investigate the weight gain of the employees of the tech company during the pandemic. The company claims that, on average, men gained more weight than women. To perform an independent-samples t-test of the hypotheses

$$H_0 : \mu_M = \mu_F \quad \text{vs.} \quad H_1 : \mu_M > \mu_F, \quad (1)$$

where μ_M and μ_F denote the expectation of the weight gain of the male and female employees, respectively, go to

Analyze → *Compare Means* → *Independent-samples T Test*

Choose **Weight** as Test Variable and **Gender** as Grouping Variable and Define Groups with the labels M and F.

► **Report the table “Independent Samples Test”**

The table also contains the results of “Levene’s test on the equality of variances”, an alternative for the F-test (see chapter 5).

- ▷ What conclusion can you draw from the p-value (= “Sig.” or “observed significance” in the table), at a 5% significance level?

The first row of the table “Independent Samples test” shows a “Sig. 2-tailed” (= 2-sided p-value).

- ▷ **Compute the one-sided p-value** for the test problem (1).

- ▷ At a significance level of 5%, **can we conclude that, on average, male employees of the company gained more weight than women** during the pandemic?

5. Copy and save the required results (**four tables**) in a Word/PDF file and add the required answers to the questions (marked by ▷). Adjust the dimensions such that it all fits in at most two pages. You can now submit to Canvas.