**Assignment: Energy Statistics**

**Teachers**

Jonathan Chambers

Imane Fouiteh

Alireza Mahmoudan

**Targets**

- Get hands-on experience on for analysing survey data related to energy & socio-economic and dwelling characteristics

**Final product**

A report not exceeding 8 pages (excluding tables) in Word or pdf. This part should be self-supporting, i.e. without links to Excel. Please explain clearly how you have made the calculations with its assumptions, and interpretation of the results. Make sure that it is well-organised and easy to check, by using different colours, naming of parameters and table captions, appropriate units and so on.

**Group size:** 2 persons per group

**Background literature**

See slides for tutorial videos.

**Deadline**

· Wednesday, 17 April 2024, 17:00

· Upload your assignments (Word/pdf and Excel document) on Moodle

· Debriefing: Friday, 19 April 2023, 8:15

**Introduction**

**Files**

Word document: Assignment

Data: SHEDS.xlsx (Swiss Household Energy Survey)

Download software (After our 2 hours class session):

SPSS: <https://plone.unige.ch/distic/pub/logiciels/spss-statistics>

**Inferential statistics:** t-test, chi-square, ANOVA test, and Linear regression

**Software to be used:** SPSS (recommended)

**Dataset**

﻿The Swiss Household Energy Demand Survey (SHEDS) was developed and designed to describe Swiss households’ energy-related behaviours comprehensively, their longitudinal changes and the existing potential for future energy demand reduction. The survey was organized in five annual waves, thus generating a rolling panel dataset of 5,000 respondents per wave.

Step 1: Get familiar with the dataset.

Step 2: Get familiar with the tools, SPSS

**Questions**

**Question 1: Descriptive Statistics. (10 points)**

Choose three socio-economic characteristics from the survey (e.g. gender, age (categorised one), dwelling type etc.) and describe each group’s electricity consumption.

**Hint:** In descriptive statistics you only describe your data (mean, median, standard deviation, min, max).

Hint: SPSS —> Analyze → Compare means → Means

* Draw a Table showing the descriptive statistics (mean, median, standard deviation, min, and max) and a bar chart to show the mean and the error bars for each socio-economic characteristic (5 points).
* Discuss the results in one paragraph for one socio-economic characteristic (5 points). What do the results imply?

**Question 2: Inferential Statistics (50 points in total)**

***2.1 Chi-square: (10 points) :***

Research question: We want to see whether there is a relationship between

* income (md\_ek) of the household and Photovoltaic ownership.
* age (age category) of the respondent and Photovoltaic ownership.

**Task:** Perform a chi-square test on PV ownership and age of the survey respondents.

**Hint:** SPSS —> Analyze → Descriptive Statistics → Cross-Tabs

**Hint2:** Statistics, tick chi-square counts: observed and expected tick percentage columns.

* Q2.1.1: Write the null hypothesis for two cases (3 points)
* Q2.1.2: Present your test results in a Table and in a text, write the values of chi-square, degree of freedom and significance of the results. (3 points)
* Q2.1.3: Interpret your results and discuss them; what do they imply? (4 points)

***2.2 t-test (10 points):***

Research question: We want to know whether electricity consumption of

(PV owners and non-owners)

(EV owners and non-owners)

(heat pump owners and non-owners) are different from each other.

Perform a t-test on the ownership and electricity consumption of the survey respondents.

**Hint:** SPSS **→** Analyze → Compare means —> Independent-Sample T-test

* Q2.2.1: Write the null hypothesis for every three cases. (3 points)
* Q2.2.2: Present your test results in a Table (3 points)
* Q2.2.3: Interpret your results (e.g. what does it mean, t-value and significance) (3 points)
* Q2.2.4**:** The t-value is negative for some of the t-tests; explain what it means. (1 point)

***2.3 ANOVA test (10 points):***

Research question: We want to know whether the electricity consumption of different household structures (seco3) is different from each other.

Information on seco3: 1: Couple with children; 2: Couple without children, 3: Single parent with one or more children; 4: Single person household; 5: Other

i) Perform a one-way ANOVA test on household type and electricity consumption of the survey respondents.

ii) Perform a post-hoc analysis (Tukey’s HSD)

**Hint:** SPSS **→** Analyze → Compare means —> One-way ANOVA

* Q2.3.1: Write the null hypothesis (3 points)
* Q2.3.2: Present your ANOVA and Tukey’s HSD test results in a Table (3 points)
* Q2.3.3: Interpret your results (e.g. what does it mean, F-value, and significance) (4 points)

**Question - 3 Regression analysis *(20 points)* :**

Research question: We want to know which factor affects the electricity consumption of a household.

i) Find 10 socio-economic variables or dwelling characteristics from the survey and run a linear regression to find the impact on electricity consumption. List the socio-economic variables. Our suggestion: age (continuous), sex (sex), dwelling\_type (dwelling type), residtype (resident type), md\_bildung (education), accom5 (floor area), md\_beruf (occupation), seco3 (household structure), md\_hhgr (number of people in the household), md\_ek (income)

ii) Prepare the variables (e.g. dummy variables etc.) to run a linear regression.

iii) Check the VIF to see if there is any collinearity between any socio-economic variables.

**Hint:** SPSS **→** Analyze → Regression —> Linear

* Q3.1: Present your regression results in a Table including the VIF values. (5 points)
* Q3.2: Interpret your regression analysis (e.g. what does it mean, F-value, t-value and significance) (3 points)
* Q3.3: State which factors are significantly impacting the electricity consumption. (5-points)
* Q3.4: Write the linear regression equation and interpret it for one variable (5-points)
* Q3.5: According to your regression model, please predict the electricity consumption of a household with the following characteristics:
	+ Age: 26
	+ Head of the household: Female
	+ Flat
	+ Degree II
	+ 90m2 floor area
	+ Retired
	+ Couple with children
	+ 3 people in the household
	+ Income of 4500 < X < 6000 CHF