Qualitative part:

1.) What are the two main ontological perspectives that were discussed in the course? Define them and explain them by using the example of the term family.

The two perspectives are objectivist and constructionist ontology

Objectivist ontology
- social phenomena confront us as external facts
- individuals are born into a pre-existing social world
- social forces and rules exert pressure on actors to conform
- e.g. culture exists independently of social actors who are socialized into its values – for example: what constitutes being a family is fixed

Constructionist (constructivist) ontology
- social phenomena and their meanings are constructed by social actors continually accomplished and revised
- researchers’ accounts of events are also constructions – many alternative interpretations
- culture does not exist independently – for example: what constitutes being a family is in constant flux and is highly subjective

2.) a) What is a research design? b) What types of research design exist (list them all)? c) How do you choose an appropriate research design? (short answers)

a.) What is a research design?
- A structure or framework to guide data collection and analysis

b.) What types of research design exist?
- experimental
- cross-sectional
- Longitudinal
- Comparative
- case study

c.) How do you choose an appropriate research design?
The decision for the appropriate research design depends on the nature of your research question(s).

3.) **What are the four main preoccupations of qualitative researchers?**

- Seeing through the eyes of those studied – taking the role of the other
- Emphasis on social process – how patterns of events unfold over time
- Flexibility and limited structure
- Concepts and theory emerge from the data without too much explicit theorizing beforehand = inductive and/or abductive

2.) **Kathy has started to work on her questionnaire and she remembered from the method class that there were some specific rules for designing good questions. What are these rules? Mention at least 5 of those.**

- avoid ambiguous terms: ‘often’, ‘regularly’, ‘frequently’, ‘have’ (people might have different frames of reference)
- avoid long questions
- Avoid two-barrelled questions, people may have different answers to each part
- don’t ask two questions in one
- avoid negative terms (‘not’, ‘never’), especially double negatives – confusing
- avoid technical terms, jargon and acronyms
- ensure that respondents have the requisite knowledge (questions make sense?)
- symmetry between closed questions and answers
- balanced between pos. and neg. responses to a question (avoid bias) *e.g., excellent, good, acceptable, poor*
- don’t rely on respondent’s memory
- include a “don’t know” option?
Task 1: Short answers

A) What is the level of measurement of the following variables from the European Social Survey (ESS)? If you were to use these four variables as independent variables in an OLS-regression, how would you code them?

1. How interested would you say you are in politics? On a scale from 1-4, where 1 is very interested and 4 is not at all interested.

Ordinal. Because the variable has less than 5 categories it is per definition not continuous and should therefore be coded into a dummy set. But students could get points for arguing that it could be used as continuous variable if the argumentation for it is good.

2. Year born.

Ratio. Should be coded into year of age when used as an explanatory variable. In SPSS: year of ESS was conducted − year born = age.

3. What is your usual gross pay before deductions for tax and insurance? [To be recorded in country’s own currency and later converted into Euros].

Ratio. Could be kept as it is.

4. Which party did you vote for in the last national election?

Nominal. Should be dummy coded in the analyses.

B) Explain briefly the following 4 terms:

1. R square

A statistical measure that determines how close the data are to the fitted regression line. Shows us how “good” the regression analysis is, explanatory power (0-1).

2. Simple random sample

With random sampling, each unit of the population has an equal probability of inclusion in the sample. Makes it possible to make inferences from the random sample to the population from which it was selected.

3. Median

Median is an alternative way to show the central tendency of a variable. Median is the middle value in the list of numbers (ranging the value on the observations from low till high).
4. Standard deviation

Standard deviation is a measure of the dispersion of a set of data from its mean. More spread=higher deviation, and vice versa.

C) What is meant by the phrase “correlation does not imply causation”? Illustrate with a relevant example.

The answer should include an explanation of what correlation and causation is (a definition), how they differ, and how they are connected. Causation refers to when an action or occurrence cause another (such as smoking causes lung cancer), while correlation is when two variables correlates with each other (such as smoking is correlated to alcoholism). Many examples are possible here, most important is that the example is relevant and that the students understand the concepts/use them correct.

Correlates here means that two variables vary together; for instance high values on variable one is associated with high values on variable two. That two things occur together does not mean that the one cause the other.

Causal relationships can be established through controlled experiments. Causality is the relation between an event (the cause) and a second event (the effect), where the second is understood as a consequence of the first.

**Task 2: OLS regression**

The questions in task 2 are all based on OLS analyses results from a study of the level of religiousness. The analysis is based on data from the Norwegian part of the European Social Survey (ESS5). The variables are as follows:

**Dependent variable:**

Regardless of whether you belong to a particular religion, how religious would you say you are?

0 Not at all religious
1
2
3
4
5
6
7
8
9
10 Very religious
Independent variables:

i) Voted last national election: Did you vote in the last [country] national election in [month/year]? 1=Yes, 0=No

ii) Born in Norway: Were you born in [country]? 1=Yes, 0=No

iii) Gender. 1=Female, 0=Male

iv) Years of full time education completed: About how many years of education have you completed, whether full-time or part-time? Please report these in full-time equivalents and include compulsory years of schooling.

v) Age of respondent: Calculated from year born.
A) Describe shortly the effect of the independent variables on the dependent variable in model 1, and the R square for model 1.

Extra points for those who see that the dependent variable is own evaluation of religiousness.

Those who voted last election see themselves as more religious than those who did not, is statistically significant at the 5-percent level (must be interpreted as dummy-set).

Those who are born in Norway see themselves as less religious than those who are not, is statistically significant at the 5-percent level (must be interpreted as dummy-set).

Females see themselves as more religious than men, is statistically significant (must be interpreted as dummy-set).
For each year of education people evaluate themselves as 0.059 less religious, is statistically significant.

Model 1 explains 7.3 per cent of the total variance in evaluation of religiousness.

B) Describe how the results changes from model 1 to model 2. Explain shortly the main difference and what effect it has on the overall results.

Here students should mention what changes the new variable, age, leads to, both in the coefficients, significance and in R square.

Major changes: Voted last election and years of education becomes insignificant. This means that there is a relationship between these variables that is not modelled properly when age is not included. Age and education has a rather simple explanation: education increases as age increases. For voted last election the connection is less explicit, but it could be that the youngest age group is not eligible to vote and that the voting variable therefore catches up some of the age-variance.

R square increases from 7.3 to 10.6 which mean that age increases the explanatory power of the model.

C) How would you improve the analysis in model 2?

Many possible solutions possible here, some suggestions are:

- Coding of variables: For instance create dummy variables to see if there are differences between groups for instance on the age variable.

- Include alternative independent variables (based on theoretical claims/hypotheses): The students could have alternative variables that they think could influence the dependent variable.

- Validity problems: For instance a more correct dependent variable? Own evaluation of level of religiosity has some validity issues.

- The students have not learned about interaction terms or squared terms, so they should not be expected to suggest this.

Students should get points here even if their suggestions are not listed here as long as it is a possible improvement of the analysis.

D) Use the results from model 2 to predict the value for a 30 year old woman, who voted last national election, who is born in Norway, and has completed 13 years of education. Show the calculation.

Level of religiousness = constant + voted + born in Norway + female + education + age

= 3.340 + 1*0.348 + 1*1.655 + 1*1.253 + 13*-0.021 + 30*0.030

= 3.340 + 0.348 – 1.655 + 1.253 -0.273 + 0.9

=3.913