MEVIT4811 exam assignment spring 2026

There are two parts in this exam. In the first part, your task will be to analyze data from PEW research center's "Wave 152 American Trends Panel", with a nationally representative panel of randomly selected US adults (N=5410). The survey was conducted in August 2024 and includes questions about two topics: (1) what American citizens think of artificial intelligence (AI); and (2) a few questions about driving and "road rage"[1].

PART 1 consists of three tasks (each with a few sub-tasks) related to the PEW dataset.

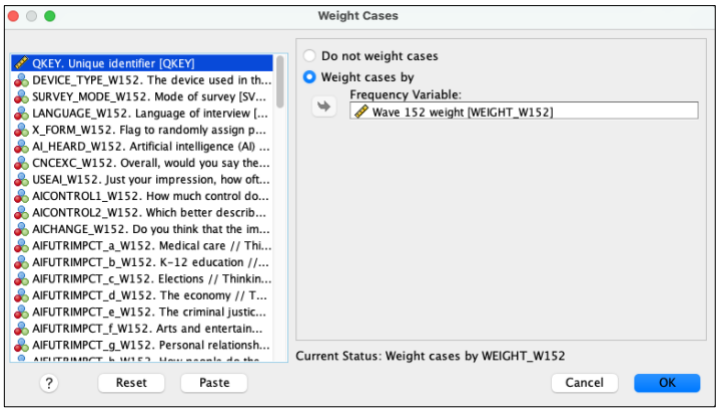
In PART 2 of the exam, we will ask you to outline a plan for a quantitative content analysis.

Before starting, we suggest you read through the entire test and then take a little time to familiarize yourself with the dataset. *Also pay close attention to what low and high values on Likert-scales indicate on the different variables.* Feel free to recode variables if you find that it makes it easier to interpret results.

There are no requirements for how many pages your exam-report needs to be.

About the PEW dataset: weight cases, the questionnaire, tips

Weight cases. Before proceeding, make sure the data are weighted. In SPSS, click on "Data" in the top menu, scroll down to "Weight Cases" and make sure the Current Status is "Weight cases by WEIGHT_W152".



If the Current Status is "Do not weight cases," find the variable "Wave 152 Weight" in bottom of the variable list and click it into the "Weight cases by" field and press OK. *Note: you will **not** get*

[1] https://www.pewresearch.org/dataset/american-trends-panel-wave-152/. If you click on the "Download dataset" on this URL, you get the sav-file and additional files about Methodology and the questionnaire. You don't need to read these to complete the exam.

a lower grade if you happen to work with unweighted data (you will only get slightly different results).

This weight has been calculated by PEW to compensate for biases in the sample compared to the population. No need to reflect on this in your exam-report.

Questionnaire. The text in the label-column in the sav-file (SPSS-file) should be sufficiently comprehensive to understand each variable.

Variables we have created. The dataset includes one variable that we have created based on the existing PEW-variables: AICONCERN_mean. This variable represents the average score derived from the items:

- AICONCERN_a_W152
- AICONCERN_b_W152
- AICONCERN_c_W152
- AICONCERN_d_W152
- AICONCERN_e_W152
- AICONCERN_f_W152
- AICONCERN_g_W152

Each of these items is measured with a 5-point ordinal scale. A reliability analysis indicated a strong correlation among these variables, supporting the assumption that they collectively reflect a common underlying construct (Cronbach's alpha= .87). We will thus treat AICONCERN_mean as a ratio-level measure of the overall level of concern respondents have toward AI. *Pay attention to what low and high scores on this new variable means* (follows from how the individual items are coded).

We have also created four dichotomized variables with values of 0 and 1:

- TRSTAIPRS_dichotomized
- Age_youngadult
- Gender_dichotomized (we here recoded "in some other way" from the original gender-variable as missing)
- Rep_rep_lean (Being/leaning Republican)

AND FINALLY: THE ACTUAL TASKS IN THIS EXAM. Best of luck to you all ☺

Overall considerations for grading exam reports:

- Precise use of statistical concepts.
- Understanding of measurement levels and ability to choose tests accordingly.
- Understanding of level of significance and probability values.
- Understanding of null hypothesis testing, types of hypotheses.

- Ability to reflect on the purpose of tests and draw meaning from the data.
- Whether students have worked with weights on or off *should not affect the grade.*

PART 1: PEW Wave 152 American Trends Panel

Task 1 includes four subtasks to help to get you warmed up

Task 1 counts for 20% of the grade.

1a) In your exam-report, include *one* frequency table with all of the following demographic variables: Age category, Gender, Education level category 2, Income tier 3-way, Party preference. Include only percentages.

Remember to set up the table in word-processing software of your choice (don't just paste a screenshot or the unedited table from SPSS).

Examiner guidelines: Excellent answers include a simple and easy to read table with correct percentage shares. Table below reports from "valid percentage", treating "refused" as missing.

	Percentage share
Age (N=5387)	
18-29	19.9
30-49	33.4
50-64	24.4
65 or older	22.3
Gender (N=5374)	
A man	48.5
A woman	50.8
In some other way	0.7
Level of education (N=5387)	
Less than high school	6.9
High school graduate	29.4
Some college, no degree	20.2
Associate degree	9.4
College graduate/some post grad	20.8
Postgraduate	13.3
Income (N=5027)	
Lower income	33.2
Middle income	48.9
Upper income	17.9
Party preference (N=5317)	
Republican	27.9
Democrat	26.2
Independent	30.4
Something else	15.4

1b) In your exam-report, report the percentage shares for each answer category on the variable TRSTAIPRS_W152 by Age category, Gender, Education level category 2, and Party summary. Briefly comment on the main patterns shown.

Again, remember to set up the table(s) in word-processing software of your (don't just paste a screenshot or the unedited table from SPSS).

Examiner guidelines: Excellent answers include a simple and easy to read table (or tables) with correct percentage shares. It's important that students make the correct choices regarding how the percentages are calculated. Excellent answers clearly and correctly comment upon patterns (e.g. emphasizing how young adult males leaning Democrat are those who are most likely to believe that AI will get to the point where they would trust it to make important decisions for them). Whether students make one table (like below) or four separate tables makes no difference on grading.

	Do you think AI will ever get to a point where you would trust it to make important decisions for you?		
	Yes, it will (%)	**No, it will not (%)**	**Not sure (%)**
Age			
18-29	17.8	59.5	22.6
30-49	14.6	62.3	23.1
50-64	12.8	65.1	22.2
65 or older	7.4	66.5	26.1
Gender			
A man	18.0	59.2	22.8
A woman	8.9	67.2	24.0
In some other way	5.3	71.1	23.7
Level of education			
Less than high school	16.2	59.3	24.5
High school graduate	10.7	63.1	26.2
Some college, no degree	9.9	69.7	20.4
Associate degree	12.3	65.4	22.3
College graduate/some post grad	15.3	62.1	22.6
Postgraduate	19.1	57.8	23
Party summary			
Rep/Lean Rep	11.2	68.5	20.3
Dem/Lean Dem	15.9	58.8	25.3

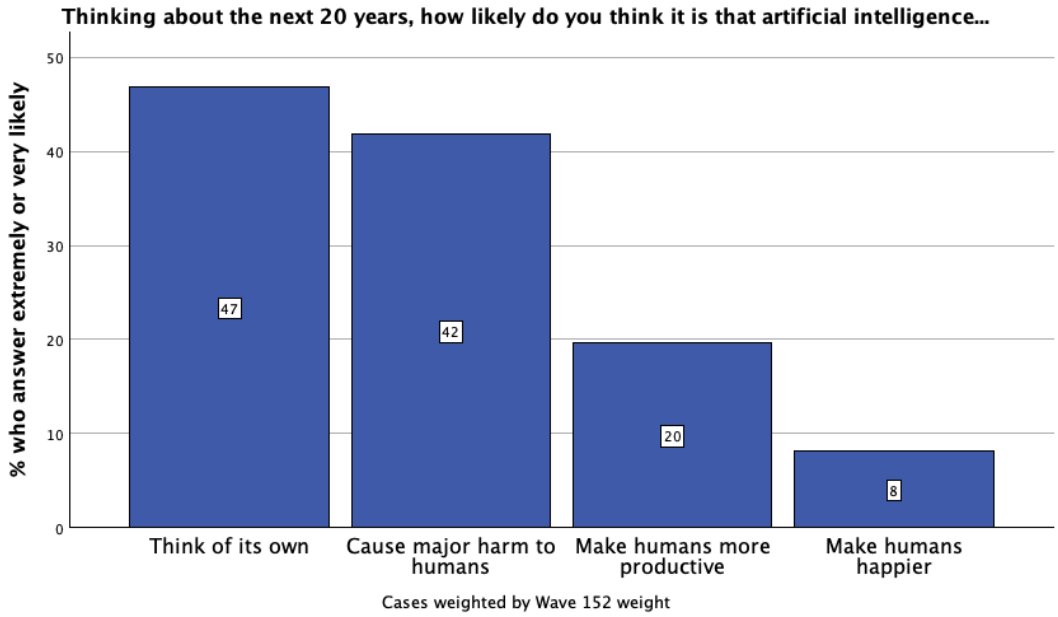
1c) Create a graph that shows how often respondents think they interact with AI (variable name: USEAI_W152) and include this in your exam-report. Explain why you chose this type of graph.

Examiner guidelines. Excellent answers explain how "USEAI_W152" is an ordinal level variable → bar chart. They include a simple and easy-to-read chart and may also have chosen to depict the Y-axis in percentages to ease interpretation for readers.

Good answers: Candidates who choose other appropriate graphs.

1d) Create *one* bar chart that shows the percentage shares who have responded "Extremely likely" or "Very likely" on the four FUTRAI-items (one bar for each item). You can create the bar chart in SPSS or Excel. Make sure to edit the graph so that it clearly conveys the results. Include the graph in your report. Briefly comment upon what the bar chart indicates.

Examiner guidelines. Excellent answers are technically correct and include the required information in the graph for readers to understand what the Y-axis and what each bar represent. The graphs should not be cluttered with redundant information (e.g., including the full text from the label-column in the sav-file for the bars). Excellent answers also comment how the bar chart indicates that US citizens overall seem to be more pessimistic than optimistic about the future consequences of AI.



Task 2 tests your ability to decide on and conduct statistical tests and interpret their results

Task 2 counts for 30% of the grade.

You will now be presented with two hypotheses and one research question. For each of these, you need to:

- Locate the appropriate variable(s) to test the hypothesis
- Explain what the correct statistical test is for the hypothesis
- Conduct the analysis
- Report and interpret the results

H1: Age is associated with the amount of control people think they have in whether artificial intelligence (AI) is used in their life. (use the age-variable "F_AGECAT")

Examiner guidelines. Candidates explain what test is appropriate and why: A chi-square test of independence. Candidates should report the results correctly in a form that explains what has been tested. For example:

A chi-square test of independence suggests a significant association between age (18-29; 30-49; 50-64; 65 or older) and the amount of control people think they have in whether AI is used in their life (a great deal; quite a bit; some; not too much; none), X2 (12, n=5356) =461.86, p< .001. Younger age is associated with larger shares thinking they have control with whether AI is used in their lives (see below cross-table).

How much control do you think you have in whether artificial intelligence (AI) is used in your life?

	A great deal	Quite a bit	Some	Not too much	None	Total
18-29	7.2 %	15.7 %	33.0 %	36.2 %	7.9 %	100 %
30-49	5.5 %	11.0 %	30.8 %	39.4 %	13.3 %	100 %
50-64	4.1 %	5.3 %	25.8 %	46.0 %	18.9 %	100 %
65+	2.2 %	3.6 %	17.2 %	44.6 %	32.4 %	100 %

Excellent answers explain the choices they make, report correctly and demonstrate an understanding of how the variables are associated.

Good answers choose the correct test and report correct results but are less elaborate in explaining the choices or interpreting results.

H2: Party preference is associated with overall level of concern toward AI.

Examiner guidelines. Candidates explain what test is appropriate and why: One-way ANOVA. Candidates should report the results correctly in a form that explains what has been tested. For example:

A one-way between-groups variance was done to examine the impact of party preference on overall level of concern toward AI. Respondents were divided into four groups according to party preference (Group 1: Republican; Group 2: Democrat; Group 3: Independent; Group 4: Something else). For the dependent variable, the minimum value is 1and the maximum value is 5. Lower scores indicate a higher level of concern toward AI.

There was a statistically significant difference, F (3, 5305) = 9.98, p< .001. The difference is very small. The effect size calculated using eta squared was .006. Post-hoc comparisons using the Bonferroni test [candidates may have used the Tukey test] indicated that the means score for Republicans (M= 2.06; SD= .79) was significantly different from Democrats (M= 2.18; SD= .76), from Independent (M= 2.20; SD= .78), and Something else (M= 2.21; SD= .85).

Excellent answers explain the choices they make, report correctly and demonstrate an understanding of how the variables are associated. They also check for homogeneity of variance.

Good answers choose the correct test and report correct results but are less elaborate in explaining the choices or interpreting results.

RQ: Does gender, younger age, being/leaning republican, and thinking AI will get to the point where respondents would trust it to make important decision for them, predict overall concern toward AI?

Tip for RQ: you might want to use the dichotomized versions of the predictor variables (with values of 0 and 1).

Examiner guidelines. Candidates explain what test is appropriate and why: Linear regression. Candidates should report the results correctly in a form that explains what has been tested. They should include a table such as the one below, and they should explain what their results mean (both the variance explained by the model, and how the different factors predict the outcome variable.

Table 1:Results from multiple regression analysis. Dependent variable: Overall concern toward AI.

	b	***SE b***	***β***	***p***
Model 1				
Constant	2.17	.021		<.001
Gender (male=0)	-.101	.022	-.065	<.001
Younger age (No=0)	.107	.028	.054	<.001
Being/leaning republican (No=0)	-.110	.022	-.070	<.001
TRSTAIPRS_dichotomized (No=0)	.438	.032	.191	<.001

Adjusted R^2=.055

Excellent answers explain the choices they make, report correctly and demonstrate an understanding of how the variables are associated. They're able to interpret results correctly (including understanding that lower score on the outcome variable means a *higher* level of concern (unless they have recoded the variable)).

Good answers choose the correct test and report correct results but are less elaborate in explaining the choices or interpreting results.

Task 3 tests your ability to work independently with the dataset and decide on appropriate statistical tests.

Task 3 counts for 30% of the grade.

Develop one research question or hypothesis based on the Pew dataset and test it using an appropriate statistical analysis in SPSS. Your analysis should examine whether one outcome variable is associated with one or more explanatory variables.

You may choose the statistical method yourself, but your choice must fit the level of measurement of the variables. *Please do not re-use one of the analyses from Task 2, though you may re-use a statistical test as long as you have a different combination of variables.* Your options here include e.g., binary logistic regression (see Pallant) in addition to tests we have covered in class. In your answer, please:

- Present the research question or hypothesis
- Explain what the correct statistical test is for this question or hypothesis
- Conduct the analysis
- Report and interpret the results

You are free to choose among all the variables, also those that have very little to do with media and communication studies.

Examiner guidelines. Candidates should present a research question or hypothesis that can be addressed with the data.

Differences between excellent, good, and fair answers relate to:

- Do candidates identify the variables and statistical tests necessary to address their research question/hypothesis? Do they understand the measurement-level of the variables?
- Does test appear to have been correctly conducted? If relevant: is choice of directional/non-directional hypothesis reflected in test? Are results correctly reported?
- Does the candidate interpret the results correctly? Does the candidate conclude what the results mean in the context of the research question and hypothesis?
- Candidates should be rewarded for choosing RQ and hypothesis that require statistical tests not included in task 2 (e.g., binary logistic regression, t-test for independent samples).
- This problem opens for candidates to show the knowledge and skills they have acquired.
- More elaborate exam-reports that have required more effort should be rewarded.

PART 2: Quantitative content analysis

Part 2 counts for 20% of the grade.

You are tasked with designing a quantitative content analysis of public social media content, focusing on one of the following topics:

- Framing of technology in AI advice or “tips” videos
- Construction of femininity in the tradwife aesthetic
- Professional news organizations adaptation to social media

Prepare a research plan that encompasses the following components:

- Formulate a research question about your chosen topic that could be answered through a quantitative content analysis. You will need to find a more specific angle on the topic.

- Identify what data material will be analyzed, including which social media platform and the unit of observation (for example one video, one post, one caption, or one comment).

- Identify **four** variables relevant to your research question and write them out as they would appear in a codebook, including the variable name, definition, and coding categories/values.

Tip: Your variables should measure "manifest" content as much as possible – that is, features that can be observed directly in the material – rather than latent qualities that require interpretation.

Examiner guidelines. Part 2 asks the candidate to demonstrate that they understand what sort of research inquiries a content analysis can answer and how this method is applied. As stated, the candidate has a lot of flexibility in how exactly they focus and formulate their inquiry. It is not necessary for their research question to have a strong theoretical justification. However, they must think through how the question might be answered. Note: It is not expected that these four variables alone would be sufficient to answer the research question that the candidate poses.

Candidates who receive a <u>good</u> grade:

- demonstrate that they understand the purpose (and limitations) of a quantitative content analysis in their formulation of the research question
- develop 4 variables with a logical connection to the research question that could plausibly be measured in a quantitative content analysis
- write a brief description of each variable and state how it is to be measured. If categorical, the categories must be included and should be exhaustive (e.g. include "Other") and mutually exclusive.

Candidates who receive <u>very good or excellent grades</u> *may*:
- develop a research question with clear relevance to existing media research
- develop a research question that seeks to find a relationship between variables
- develop a detailed research question that focuses on a specific aspect of their theme
- write variables that operationalize theoretical concepts introduced in the research question
- demonstrate sophisticated thinking in their development of their variables – e.g. with precise definitions and well-thought-out levels of measurement
- include detailed coder instructions for how to measure the variables in the hypothetical data material