※ 注意：請用 2B 鉛筆作答於答案卡，並先詳閱答案卡上之「畫記說明」。

單選題，請使用2B鉛筆作答於答案卡，每題5分

Choose the most suitable answer and mark it on your answer sheet. Each question carries 5 points.

Questions 1 to 4 are based on the following information:

A research study by Dai, Milkman, and Riis, published in 2014 at *Management Science*, suggests that people are more likely to tackle their goals immediately following salient temporal landmarks. It is a psychological phenomenon where people feel motivated to pursue goals and make positive changes at temporal landmarks, such as the beginning of a new year, month, week. To verify the hypotheses, the authors obtained data from "Google Insights for Search" to download the daily number of Google web searches that include a given search term. They downloaded data on the daily number of Google searches in the United States for the term "diet" from January 1, 2004, to June 30, 2012 (a time period including 3,104 days). If people are more likely to tackle their goals immediately following salient temporal landmarks, then the search behavior of the term "diet" may change as the time goes.

Daily search volume data provided by Google Insights for Search is normalized to fall between 0 and 100. The authors examine whether people are more interested in dieting following temporal landmarks using ordinary least squares (OLS) regression analyses. The predictor variables in the OLS regressions include measures of a given day's distance from the beginning of the week (from 1 = Monday to 7 = Sunday), month (days elapsed since the beginning of the current month with min = 1 and max = 31), and year (from 1 = January to 12 = December).

To further exclude the possibility that the findings can be attributed to general patterns of Internet search over time, the authors compare searches for the term "diet" with searches for two popular search terms, "news" and "weather", which do not relate to aspirational behaviors. The regression results (regression coefficients and standard errors in parentheses) are shown in the following table.

Google search term	**Diet**	**News**	**Weather**
Days since the start of the week	-1.63^{***}	-2.09^{***}	0.72^{***}
	(0.08)	(0.11)	(0.17)
Days since the start of the month	-0.09^{***}	-0.05^{*}	0.09
	(0.02)	(0.02)	(0.04)
Months since the start of the year	-3.81^{***}	-0.05	0.93
	(0.42)	(0.45)	(0.83)
Other Control Variables	Yes	Yes	Yes
Observations	3104	3104	3104
R2	0.62	0.81	0.53

$^{*}p < .05$; $^{**}p < .001$; $^{***}p < .0001$

Data source: Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science*, *60*(10), 2563-2582.

Please answer questions 1 to 4 based on the information above.

見背面

1. The study uses Google search data and comparison search terms to test a behavioral hypothesis about goal initiation. Based on the research design and background information provided, which of the following statements is NOT correct?
   A) Dieting commonly begins on psychologically salient temporal landmarks such as Mondays, the beginning of a new month, or the start of a new year.
   B) "Diet" is an appropriate example of goal pursuit because it reflects a common, aspirational, self-regulatory behavior that individuals frequently intend to initiate at meaningful temporal landmarks.
   C) The authors include search terms such as "news" and "weather" in their regression analysis primarily to increase the sample size and improve statistical power.
   D) Google search activity for the term "diet" serves as a proxy for individuals' interest in initiating dieting behavior, even though it may not directly measure actual dieting outcomes.
   E) Comparing "diet" searches with non-aspirational terms helps rule out the possibility that observed temporal patterns are driven by general Internet search behavior rather than goal-related motivation.

2. Based on the regression results reported in the table, and interpreting all statements as referring to model-predicted search volume while holding other regressors constant, which of the following statements is NOT correct?
   A) The model implies that the predicted difference in search volume for "diet" between Monday and Tuesday is the same as the predicted difference between Saturday and Sunday.
   B) The model predicts that the search volume for "diet" tends to decline as the month progresses from its beginning toward its end.
   C) On average, the model predicts that daily search volume for "diet" decreases as the week progresses from Monday to Sunday.
   D) The regression results indicate a statistically significant difference in average "diet" search volume between the beginning of the year and the middle of the year.
   E) After controlling for month-of-year effects, temporal landmarks such as the start of a new week have no additional association with goal-related search behavior.

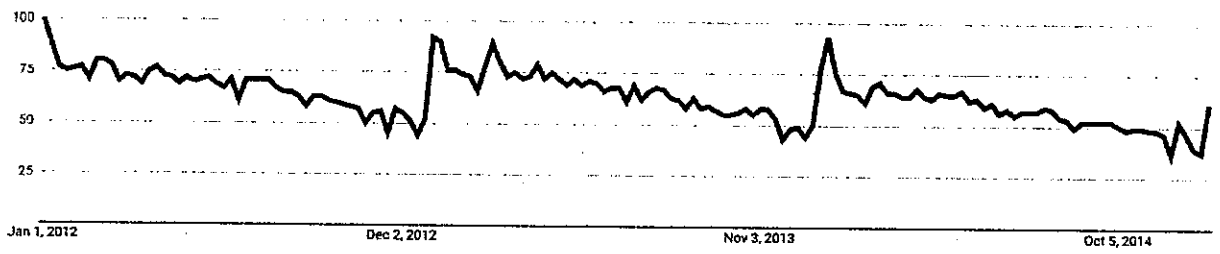
3. Based on the regression results reported in the table, and holding the other regressors constant, how many of the following six statements are supported by the sign and statistical significance of the estimated coefficients for the control search terms ("news" and "weather")?
   - The predicted search volume for "news" significantly decreases as the week progresses from Monday to Sunday.
   - The predicted search volume for "news" significantly decreases as the month progresses from the beginning to the end of the month.
   - The predicted search volume for "news" significantly decreases as the year progresses from the beginning to the end of the year.
   - The predicted search volume for "weather" significantly decreases as the week progresses from Monday to Sunday.
   - The predicted search volume for "weather" significantly decreases as the month progresses from the beginning to the end of the month.
   - The predicted search volume for "weather" significantly decreases as the year progresses from the beginning to the end of the year.

   A) None of the statements are supported.
   B) Only one of the statements is supported.
   C) Two of the statements are supported.
   D) Three of the statements are supported.
   E) More than three statements are supported.

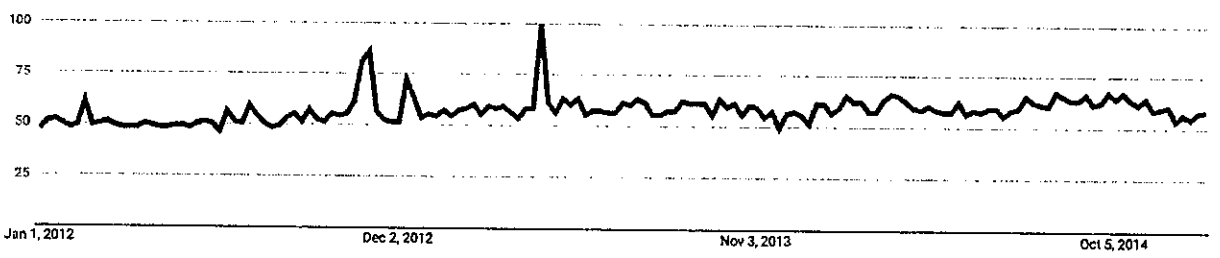
接次頁

4. Based on the regression results reported in the table and the background information provided, which of the following charts is most likely to represent the time series pattern of the daily Google search volume for the term "diet" over the period January 1, 2012 to December 31, 2014?

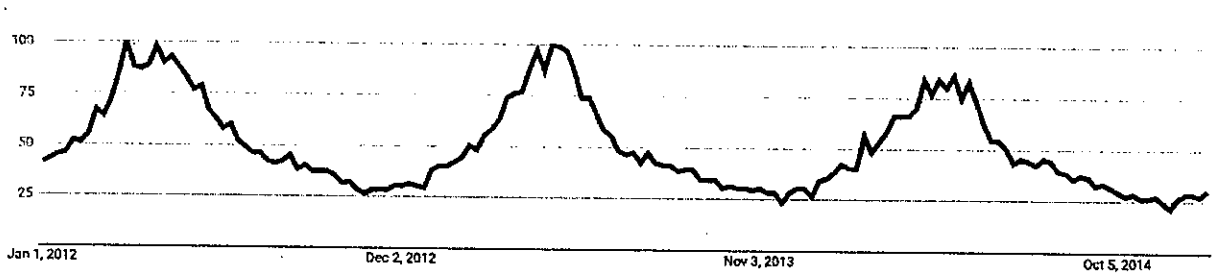
A)



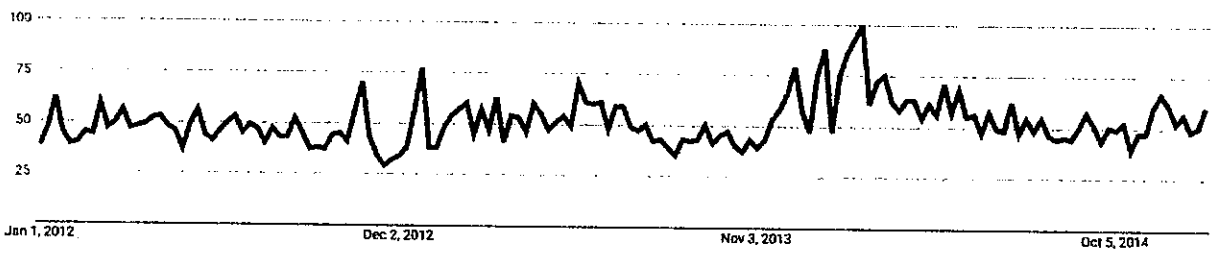
B)



C)



D)



E) It is not possible to rule out any of the charts, because the regression analysis is based on data only through June 30, 2012, and no evidence is available on search behavior beyond that period.

Questions 5 to 7 are based on the following information:

A research study by Acemoglu (2022), published in *Economica*, examines how the values that families emphasize for their children vary with family income. The analysis uses data from the World Values Survey (WVS), a series of nationally representative surveys covering 113 countries and approximately 432,463 respondents over the period 1981–2018.

Two outcome variables—independence and obedience—are constructed as binary indicators that equal 1 if the respondent mentions the corresponding characteristic when answering the survey question: "Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider especially important?"

Family income is measured using a self-reported ordinal scale based on the survey question: "Here is a scale of incomes. We would like to know in what group your household is, counting all wages, salaries, pensions, and other incomes that come in. The scale goes from 1 to 10, representing the lowest and highest values."

The figure below shows unconditional averages of the probability that respondents emphasize independence or obedience as important child qualities across family income categories. Error bars correspond to two–standard error confidence intervals.

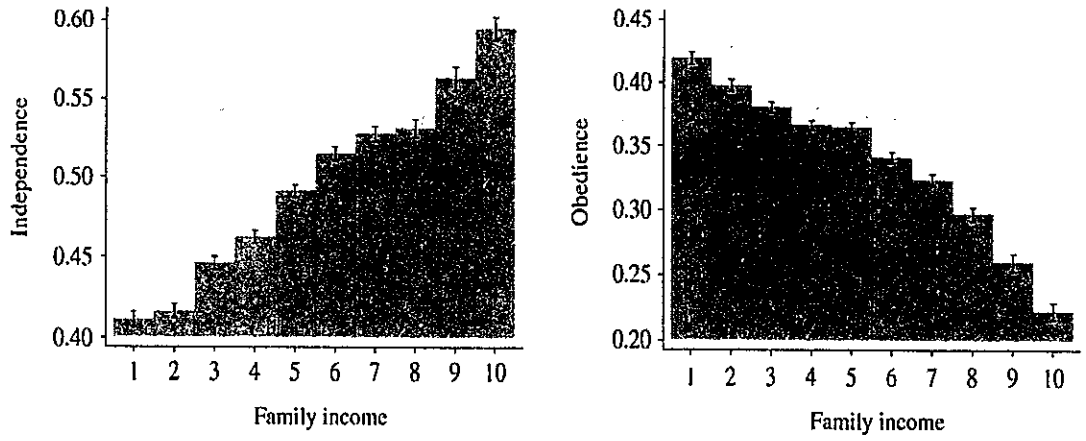


Figure source: Acemoglu, D. (2022). Obedience in the labour market and social mobility: a socioeconomic approach. *Economica*, *89*, S2-S37.

Please answer questions 5 to 7 based on the information above.

5. Based on the background information and the figure described above, which of the following statements is correct?
   A) A larger proportion of respondents report that obedience is especially important for children than report that independence is especially important.
   B) Because the data are nationally representative and span many countries and years, the observed relationship between family income and parental values can be interpreted as causal.
   C) The figure provides evidence of an association between family income and the probability that respondents emphasize independence or obedience as important child qualities.
   D) Each respondent is restricted to selecting only one child quality in the survey.
   E) None of the statements above can be concluded based on the information provided.

接次頁

6. Suppose the researcher conducts a correlation analysis between each outcome variable (independence and obedience) and family income. Based on the patterns shown in the figure, which of the following results is most plausible?
   A) The correlation coefficient between independence and family income is significantly positive, and the correlation coefficient between obedience and family income is significantly positive.
   B) The correlation coefficient between independence and family income is significantly positive, and the correlation coefficient between obedience and family income is significantly negative.
   C) The correlation coefficient between independence and family income is significantly negative, and the correlation coefficient between obedience and family income is significantly positive.
   D) The correlation coefficient between independence and family income is not significantly different from zero, while the correlation coefficient between obedience and family income is significantly negative.
   E) A correlation analysis cannot be conducted because the outcome variables are binary indicators.

7. Suppose the researcher estimates separate logistic regression models for each outcome variable—independence and obedience—using family income as the explanatory variable. Based on the patterns shown in the figure, which of the following results is most plausible?
   A) The estimated coefficient on family income is positive and statistically significant in both the independence and obedience regressions.
   B) The estimated coefficient on family income is positive and statistically significant in the independence regression, and negative and statistically significant in the obedience regression.
   C) The estimated coefficient on family income is negative and statistically significant in the independence regression, and positive and statistically significant in the obedience regression.
   D) The estimated coefficient on family income is negative and statistically significant in both the independence and obedience regressions.
   E) Logistic regression is inappropriate in this context because family income is measured on an ordinal scale.

Questions 8 to 10 are based on the following information:

A research study by Mookerjee, Cornil, and Hoegg (2021), published in the *Journal of Marketing*, investigates whether explicitly highlighting cosmetic imperfections in produce through "ugly" labeling (e.g., labeling cosmetically imperfect cucumbers as "Ugly Cucumbers") can increase consumers' likelihood of purchasing aesthetically unattractive produce.

In one field experiment conducted at a farmers' market, the researchers sold pairs of produce items consisting of one aesthetically unattractive version and one aesthetically attractive version of the same product (e.g., potatoes, carrots, tomatoes). Each version was displayed in separate baskets with labels attached. The labeling of the unattractive produce was experimentally manipulated: in the "ugly" label condition, the unattractive produce was explicitly labeled as "Ugly" (e.g., "Ugly Potatoes," "Ugly Carrots," "Ugly Tomatoes"), whereas in the control condition, it was labeled without the "ugly" descriptor (e.g., "Potatoes," "Carrots," "Tomatoes"). Across both conditions, the attractive produce was always labeled simply by the product name (e.g., "Potatoes," "Carrots," "Tomatoes"). In all conditions, the unattractive produce was sold at a 25% discount relative to the attractive produce.

The authors alternated the labels used for the unattractive produce every hour, which helped approximate random assignment of consumers to the "ugly" label and control conditions based on the timing of their visit. Over four days, 938 individuals stopped at the stand, of whom 259 interacted with the seller, and 113 made at least one purchase.

見背面

The following figure presents mean spending by buyers (in local currency CAD) on unattractive and attractive produce, separately for the control label condition and the "ugly" label condition. In the control condition, unattractive produce is not labeled in any specific way, whereas in the "ugly" label condition, unattractive produce is explicitly labeled as "ugly." Error bars represent ±1 standard error. Statistical significance of differences between label conditions is indicated by symbols above the bars (†$p < 0.10$; *$p < 0.05$).

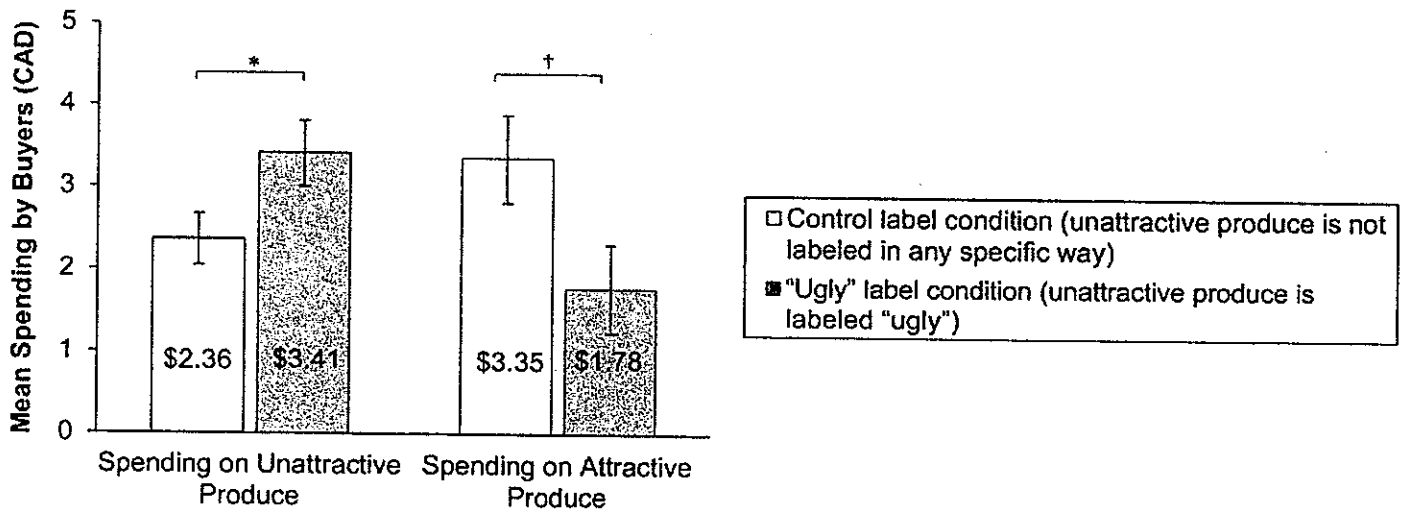


Figure source: Mookerjee, S., Cornil, Y., & Hoegg, J. (2021). From waste to taste: How "ugly" labels can increase purchase of unattractive produce. *Journal of Marketing*, *85*(3), 62-77.

Please answer questions 8 to 10 based on the information above.

8. In the control condition, 62.5% of buyers purchased unattractive produce and 56.0% purchased attractive produce. In the "ugly" label condition, 81.6% of buyers purchased unattractive produce and 26.5% purchased attractive produce. (Percentages do not sum to 100% because some buyers purchased both types of produce.) Which of the following statement is correct?
   A) A logistic regression can be used to examine whether labeling unattractive produce as "ugly" (versus the control label) affects buyers' likelihood of purchasing unattractive produce.
   B) A linear regression can be used to examine whether labeling unattractive produce as "ugly" (versus the control label) affects buyers' likelihood of purchasing unattractive produce.
   C) The reported percentages alone are sufficient to conclude that labeling unattractive produce as "ugly" caused a change in purchasing behavior.
   D) Labeling unattractive produce as "ugly" (versus the control label) significantly increased buyers' likelihood of purchasing attractive produce.
   E) Labeling unattractive produce as "ugly" (versus the control label) significantly decreased buyers' likelihood of purchasing unattractive produce.

接次頁

9. Based on the figure showing spending as the dependent variable, which of the following analytical approaches is most appropriate for analyzing the data?
   A) An independent-samples T test comparing average spending between the "ugly" label condition and the control condition.
   B) A paired t test comparing spending on unattractive produce and attractive produce, ignoring label condition.
   C) A linear regression of spending on a single dummy variable indicating whether the produce is labeled "ugly," without accounting for produce appearance.
   D) A mixed-effects (or mixed-design) regression/ANOVA of spending, with label condition ("ugly" vs. control) as a between-subjects factor and produce appearance (unattractive vs. attractive) as a within-subject factor.
   E) No formal statistical analysis is appropriate because some buyers purchased both attractive and unattractive produce.

10. Based on the figure showing mean spending by buyers, which of the following statements is correct? (Interpret dollar amounts as sample means, and statistical significance as indicated in the figure.)
   A) In the control condition, buyers spent on average $2.36 on unattractive produce and $3.41 on attractive produce.
   B) In the "ugly" label condition, buyers spent on average $3.35 on unattractive produce and $1.78 on attractive produce.
   C) The figure indicates that the main effect of label condition ("ugly" vs. control) on total spending is not statistically significant.
   D) The figure indicates a statistically significant interaction effect between label condition ("ugly" vs. control) and produce appearance (attractive vs. unattractive).
   E) All of the statements above are correct.

11. A scatter diagram reveals a strong positive linear relationship between two variables. Which of the following numerical techniques will not give us more detailed information about this relationship?
   A) Coefficient of correlation
   B) Coefficient of variation
   C) Coefficient of determination
   D) All of these choices help us describe this relationship
   E) None of these choices helps us describe this relationship

12. Which of the following is true regarding the sampling distribution of the mean for a large sample size? Assume the population distribution is not normal.
   A) It has the same shape, mean, and standard deviation as the population.
   B) It has the same mean and standard deviation as the population, but a different shape.
   C) It has the same shape and mean as the population, but a different standard deviation.
   D) It has the same mean as the population, but a different shape and standard deviation.
   E) It has a different shape, mean, and standard deviation compared to the population.

13. In a one-way ANOVA, error variability is computed as the sum of the squared errors, SSE, for all values of the response variable. This variability is the:
   A) total variation
   B) within-treatment variation
   C) covariation
   D) between-treatment variation
   E) None of the choices above

見背面

14. If we are testing for the difference between the means of two independent populations with equal variances, samples of $n_1 = 15$ and $n_2 = 15$ are taken, then the number of degrees of freedom is equal to
A) 13
B) 14
C) 15
D) 28
E) 29

15. A candidate can earn the Chartered Financial Analyst (CFA) designation only after passing three annual examinations: CFA I, II, and III. Candidates who pass an examination are eligible to take the examination for the next level in the following year. The pass rates for levels I, II, and III are 57%, 73%, and 85%, respectively. Suppose that 3,000 candidates take the level I examination, 2,500 take the level II examination, and 2,000 take the level III examination.
Now, consider a randomly selected candidate who is going to take an examination. If the candidate manages to pass the examination, what is the probability that he or she is now effectively a CFA already? Choose the closest answer.
A) 0.220
B) 0.228
C) 0.243
D) 0.325
E) 0.348

16. A recent election study showed that 45% of registered American voters are Democrats, 42% are Republicans, and 13% are others. If two registered voters are selected at random, provided that they are either Democrats or Republicans, what is the probability that they have the same party affiliation? Choose the closest answer.
A) 0.495
B) 0.498
C) 0.501
D) 0.504
E) 0.507

17. What statistical test do we use to test the validity of a regression model?
A) F-test
B) z-test
C) t-test
D) chi-square test
E) binomial test

18. Chlorofluorocarbons (CFCs) damage the ozone layer. A city is planning to ban the use of CFCs in vehicle air conditioners. Suppose there are 5.8 million vehicles in the city, and you guess 20% of them use CFCs in air conditioners. How many vehicles will you need to sample to estimate the number of vehicles using CFCs with a margin of error of 0.3 million vehicles, with 95% confidence? Choose the closest answer.
A) 200
B) 210
C) 220
D) 230

接次頁

E) 300

19. You are fitting a multiple linear regression model. Which of the following is always true about the adjusted R^2 of this model?
A) It will increase when a predictor is added.
B) It will increase when a predictor is added and if the added predictor's p-value is < 0.05.
C) It will increase when a predictor is added and if the added predictor is not correlated with all existing predictors.
D) It will increase whenever R2 increases.
E) It increases if the t-statistic for the added predictor satisfies $|t| > 1$.

20. The difference between a sample mean and the population mean is called:
A) selection bias
B) survival error
C) sampling error
D) non-sampling error
E) nonresponse error

試題隨卷繳回