	University of Amsterdam Faculty of Economics and Business Department Economics and Econometrics	**Example Midterm exam** **Quantitative Data Analysis 1** Time: 90 minutes	

This exam uses a decimal point.
Give your **final answers** on the separate answer form. You can only give your final answers, so no calculations, motivations, explanations, etc.
There are 10 questions. Every question is 1 point.

The questions:

Question 1

In a data set of a continuous quantitative variable a certain observation has z-value -1.52. The mean of this variable is 172 and the standard deviation is 8. Find the value of this observation.
(***2 decimals***)

Question 2

Given is the sample (n=12): 3 6 12 16 17 17 19 22 25 30 33 37
If you were to make a box plot of this sample, then what would be the upper fence (which is the boundary line for identifying outliers)? (***3 decimals***)

Question 3

The relation between the number of courses, UvA students passed in year 1 (Y) and the attendance rate (in a percentage) during lectures (*X*) is described by the regression line:
$\hat{Y} = 2 + 0.08X$. Pearson's coefficient of correlation has been measured: r = 0.432.
For each course a student passes, 6 study points are earned.
Give the slope of the regression line, when the number of study points (*W*) is predicted from the attendance rate. (***2 decimals***)

Question 4

32% of the employees of a company did an anti-stress training. Of the employees who did an anti-stress training, 11% is having stress, while of the employees who did not do anti-stress training, 24% is having stress. Find the probability that an employee with stress did do anti-stress training.
(***4 decimals***)

Question 5

In a company the number of products sold on a day are X for product A and Y for product B. X and Y have the following distribution:

X \ Y	1	2	total
0	0.3	0.2	0.5
1	0.1	0.4	0.5
total	0.4	0.6	1.0

From this table we get E(X)=0.5 and E(Y)=1.6.
Calculate the standard deviation of X. (***2 decimals***)

Question 6

See previous question. Calculate the covariance of X and Y. (***2 decimals***)

Question 7

The number of regular bikes (X) and the number of e-bikes (Y) sold by a dealer in a month have E(X)=9, σ(X)=2, E(Y)=13, σ(Y)=3, COV(X,Y)=1. The monthly profit is W=0.9X+1.5Y–24. What is the standard deviation of monthly profit? (***3 decimals***)

Question 8

The production of an employee is considered to be normally distributed. The expected value μ is equal to 17 and the standard deviation σ is unknown. But it is known that 10% of the production is lower than 12. Find the value for the standard deviation σ.
(***2 decimals***)

Question 9

The income (in €1000) of a random household has a non-normal distribution with mean 42 and standard deviation 9. A random sample of 100 households is taken. Find the probability that the sample mean will be below 40. (***4 decimals***)

Question 10

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Hourly gross wage in national currency	200	4.57	83.82	15.2700	10.18788
Valid N (listwise)	200				

Find the lower bound of the 99% confidence interval of the population mean of the Hourly gross wage. **(*2 decimals*)**

Example Midterm exam
Quantitative Data Analysis 1
Time: 90 minutes

Solutions

At the exam **you can only give your final answer** (usually just one number), without any further calculations, motivations, explanations, etc. So make sure you do not make any calculation error!

<table>
<tr><th></th><th>ANSWER</th><th>Explanation (cannot be given at the exam!)</th></tr>
<tr><td>Question 1.</td><td>159.84</td><td>$172 - 1.52*8 = 159.84$</td></tr>
<tr><td>Question 2.</td><td>52.375</td><td>L25 = 3.25, Q1=12+0.25(16-12)=13
L75 = 9.75, Q3=25+0.75(30-25)=28.75
Upper Fence =28.75+1.5(28.75-13)=52.375</td></tr>
<tr><td>Question 3.</td><td>0.48</td><td>$\hat{W} = 6\hat{Y} = 12 + 0.48*X$</td></tr>
<tr><td>Question 4.</td><td>0.1774</td><td><table><tr><td></td><td>Has stress</td><td>Has no stress</td><td>Total</td></tr><tr><td>Did training</td><td>0.11*0.32=0.0352</td><td>0.32-0.0352=0.2848</td><td>0.32</td></tr><tr><td>No training</td><td>0.24*0.68=0.1632</td><td>0.68-0.1632=0.5168</td><td>1-0.32=0.68</td></tr><tr><td>Total</td><td>0.0352+0.1632=0.1984</td><td>1-0.1984=0.8016</td><td>1</td></tr></table>P(did training | has stress) = 0.0352 / 0.1984 = 0.1774</td></tr>
<tr><td>Question 5.</td><td>0.50</td><td>$V(X) = E(X^2) - (\mu_x)^2 = 1^2\ 0.5 - 0.5^2 = 0.25$
$\sigma(X)=0.5$</td></tr>
<tr><td>Question 6.</td><td>0.10</td><td>$COV(X,Y) = E(XY) - \mu_x\mu_y = 1*1*0.1 + 1*2*0.4 - 0.5*1.6 = 0.1$</td></tr>
<tr><td>Question 7.</td><td>5.118</td><td>$V(W) = 0.9^2\,V(X) + 1.5^2\,V(Y) + 2*0.9*1.5*COV(X,Y) =$
$= 0.9^2\,2^2 + 1.5^2\,3^2 + 2*0.9*1.5*1 = 26.19$
$\sigma(W) = 5.118$</td></tr>
<tr><td>Question 8.</td><td>3.91</td><td>The z-value for 10% left-tailed area is -1.28
So $\mu - 1.28*\sigma = 12$.
Substitute $\mu = 17$. Then you get: $17 - 1.28\sigma = 12$.
So $1.28\sigma = 17 - 12 = 5$ and $\sigma = 5/1.28 = 3.90625$</td></tr>
<tr><td>Question 9.</td><td>0.0132</td><td>$P(\bar{X} < 40) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{40-42}{9/\sqrt{100}}\right) = P(Z < -2.22) = 1 - 0.9868 = 0.0132$</td></tr>
<tr><td>Question 10.</td><td>13.40</td><td>Df=199≈200, $t_{.005}$ =2.601, so lower bound = $15.27 - 2.601*10.18788/\sqrt{200}$ = 13.39626</td></tr>
</table>