Problem 1

1.

.

0,4028

2.

0,38158

Problem 2

1.

The expected number of wash trades per person among persons involved in wash trading in Denmark can be calculated as follows:

(1*0,83)+(2*0,1)+(3*0,04)+(4*0,02)+(5*0,01) = 1,280

The expected number of wash trades per person equals 1,280.

Problem 3

1.

Five number-number summary for boys: 0,11111 – 0,5 – 0,75 – 0,875 - 1

Five number-number summary for girls: 0,11111 – 0,375 – 0,61111 – 0,75 – 1

For boys, the middle 50% of observations are between 0,5 and 0,875. For girls, the middle 50% of observations is between 0,375 and 0,75. In the boy distribution, the middle 50% is noticeably closer to the maximum (1) than in the girl distribution. In other words, more observed boys than observed girls believe that a person of the same sex is “really, really smart”.

Problem 4

1.

I will use the probability proportion test. The hypotheses are as following:

H0: p = 0,946

HA: p ≠ 0,946

Using software, the test statistic equals -12,323071. This is approximately standard normally distributed under the null hypothesis giving us a p-value smaller than 0,0000. As a result, we reject the null hypothesis and conclude that the probability of homeless older adults being vaccinated is smaller (0,780142) than that of the general Danish population (0,946).

Problem 5

1.

We are given two proportions, therefore I will conduct a test comparing the two proportions. The hypothesis is as following:

H0: p1 = p2

HA: p1 ≠ p2

The proportion of women lying on their dating profile is 0,4250. The proportion of men lying on their dating profile is 0,5500. The difference between the probabilities is 0,1250.

Using software, the 95% confidence interval for the difference is:

]-0,34234 ; 0,09234[

As a result, we are 95% confident that the difference between the proportions of men and women lying about their height on their online dating profiles is between -0,34234 and 0,09234. As our predicted difference in proportion includes 0, we can not assume a difference in proportion between men and women lying about their heights on their online dating profiles.

Problem 6

1.

I will use the two sample t-test assuming unequal standard deviations. The hypotheses are as stated:

H0: $\mu 1 = \mu 2$

HA: $\mu 1 \neq \mu 2$

Using software, the 99% confidence interval for the difference between the two groups is:

]9,50789 ; 13,49211[

using a t-quantile with 137,066 degrees of freedom. Based on this, we can say with 99% confidence that the drill & blast group on average experiences approximately between 9,5 and 13,5 more measured dust exposure than the outdoor concrete group.

Problem 7

1.

I will use the chi squared test. The hypotheses are stated below:

H0: The two variables are independent.

HA: The two variables are not independent.

Using software, the x2-test statistic equals 9,0825. It is x2-distributed with 2 degrees of freedom under the null hypothesis of no difference. The p-value is 0,01066. As a result, we reject the null hypothesis and conclude that there is a difference between gender and amount of oxygen required. Looking at either the mosaic plot or the estimated probabilities, it is evident that a higher proportion of females than males required no oxygen, and a higher proportion of males than females required ventilation.

Problem 8

1.

I will use the one-way anova. The hypotheses are as following:

H0: No difference between the group means.

HA: At least two of the population means are not equal.

Using JMP, the test statistic equals 7,3571 and is F-distributed under the null hypothesis with 2 and 237 degrees of freedom, which gives us a p-value of 0,0008. As a result, we reject the null hypothesis and conclude that the degree to which a child thinks that the person of the same sex is "really, really smart" depends on the age of the child. Using Tukey's HSD, it is clear that the mean of 5-year olds (0,71892) is significantly higher than 6-year olds (0,58750) and 7-year olds (0,60399). In other words, the 5-year old group has a higher tendency than the other two groups to have the stated belief.

2.

I will use the two-way anova. The hypotheses are as following:

H0: No interaction.

HA: Interaction.

Using JMP, the test statistic for testing the hypothesis of no interaction equals 3,0386. It is F-distributed with 2 and 234 degrees of freedom under the null. The p-value is 0,0498. As a result, we can (barely) reject the null hypothesis and we conclude that the effect of age on the degree to which a child thinks that the person of the same sex is "really, really smart" is likely to depend on sex.