Problem 1

P(T) = probability of being a twin = 0,1

P(L) = probability of being left-handed = 0,034

P(L given T)= 0,21

1. P(L and T) = P(T)*P(L given T) = 0,1 * 0,21 = 0,021 = 2,1%
   The probability that a person from Belgium is left-handed and a twin 2,1%.

2. P (T given L) = P(L and T)/ P(L) = 0,021 / 0,034 = 0,618 = 61,8%
   The probability that the left-handed person you just met is a twin is 61,8%.

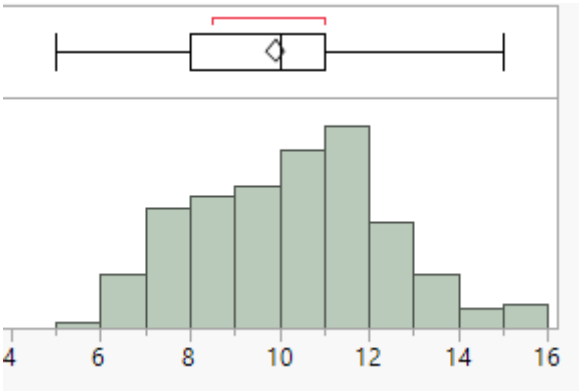
Problem 2

1. P (X<600) = P($\frac{X-674}{42} < \frac{600-674}{42}$)= P(Z<-1,7619) = 0,0390 calculated with excel

The probability of the battery of a new iPad lasting less than 10 hours is 3,9%.

Problem 3

1. The histogram:



The shape of this histogram is unimodal and left-skewed. The five-number summary for the shoe sizes is : [5, 8, 10, 11, 15].

Problem 4

1. We are for this test comparing proportions. So we will do a significance test for the difference between the proportions:

- Assumptions: We expect that the data for this test has been collected by randomization and we have a categorical response variable observed in each group. The two groups are independent.

And $n_1\hat{p}_1 \geq 5, n_1(1-\hat{p}_1) \geq 5, n_2\hat{p}_2 \geq 5, n_2(1-\hat{p}_2) \geq 5$ needs to be true, but it isn't because n_2=16 and $(1-\hat{p}_2)$ = 1/16 and so $n_2*(1-\hat{p}_2)$= 1
So this test can't be trusted.

- H_0 : $p_1=p_2$
  H_a : $p_1 \neq p_2$
- We use these formulas :

$$se = \sqrt{\hat{p}(1-\hat{p}) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = 0{,}1211$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{se_0} = -2{,}0398$$

- We get a z test statistic of -2,0398 and a p-value of 0,0414.
- Thus we can reject the null hypothesis, because the p-value is lower than the significance level of 0,05 and therefore we can say with 95% certainty that there is a difference between the two groups.
  p_1 hat = 58/84 = 0,6905
  p_2 hat = 15/16 = 0,9375
  The estimate for the difference is : -0,2470
  We can conclude that for people suffering from PLDM, more left-handed suffer from double-sided PLDM tahn right-handed (by calculating the difference between the proportions or looking at the z-statstic which is negative).

Problem 5

1. In this study 93 children out of 135 thinks that a bearded man is tronger than a clean-shaven one. So we estimate the probability of a child thinking that a bearded man is stronger than a clean-shaven one as 93/135 which is approximately 0,6889. A 95% confidence interval for this probability is [0,6108 ; 0,7670] with the formulas :

$$sample\ proportion \pm 1.96 * standard\ deviation \leftrightarrow \hat{p} \pm z * se$$

$$se = \sqrt{\frac{\hat{p}*(1-\hat{p})}{n}} = \sqrt{\frac{0{,}6889*0{,}3111}{135}}$$

2. The children who have been asked the questions are the same children for both questions therefore we will do a McNemar's test, which tests whether the proportions are the same for bearded men perceivedd as stronger and bearded men perceived as more dad-like:

	Test 2 positive	Test 2 negative
Test 1 positive	*a*	*b*
Test 1 negative	*c*	*d*

The assumption for this test is that the cells b and c sum up to more than 30 or that the sum is equal to 30. In this casse the sum is 69 so this assumption is met. The z-statistic is 4,4543 for this test and the p-value is 0,00001. Therefore we can reject the null and say that there is a significant difference between the

proportion of children who sees a bearded man as the stronger and as more dad-like. 56 out of 135 children think that a bearded man is more dad-like, which is a proportion of around 0,4148 whereas as we calculted in the question 1, 0,6889 believe that a bearded man is stronger.
So we can conclude that the probability of a child perceiving a bearded man as stronger is higher than the probability of a child perceiving a bearded man as dad-like.

Problem 6

1. We have to see if the difference between these means is actually significantly different or not. We are allowed to assume that the standard deviations are equal in this case because s_1/s_2 = 709,7/593,9 = 1,1950 approximately which is below 2.
So we do a confidence interval for the difference using the following formulas:

$$se = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} * \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 238{,}9409$$

$$df = n_1 + n_2 - 2 = 28$$

And we get : [-147,4482 ; 831,4482]
This 95% confidence interval contains 0, so we can't be sure that there actually is a difference between the annual contributions of the two account types.

Problem 7

1. To test whether the effect of height differs between sexes in this model, we have to look at the interaction between the two. The t-test for the interaction is -0,51, which is t-distributed with 403 degrees of freedom and the p-value is 0,6095 which is above the significance level of 0,05 and means that we can't reject the null hypothesis and therefore can't say that the effect of height differs between sexes.

Indicator Function Parameterization

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-13,81311	1,451839	-9,51	<,0001*
height	0,3530954	0,020393	17,31	<,0001*
sex[F]	-0,990937	0,133268	-7,44	<,0001*
(height-68,4238)*sex[F]	-0,016106	0,031503	-0,51	0,6095

With this model the shoe size of a male who is 70 inches tall is:

-13,8131 + 0,3531*70 = 10,9039

We then get the 95% prediction interval for this shoe size from JMP :

[9,0322;12,7749]

2. To test whether shoe size depends on height when sex is taken into account, we use the test statistic given by JMP which is 22,3, t-distributed with 404 degrees of freedom and which gives a p-value smaller than 0,0001, which means that the parameter is significant. So the shoe size depends on height when sex is taken into account.

Indicator Function Parameterization

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%
Intercept	-13,33308	1,106341	-12,05	<,0001*	-15,50798	-11,15817
height	0,3463461	0,015529	22,30	<,0001*	0,3158176	0,3768745
sex[F]	-0,979445	0,131239	-7,46	<,0001*	-1,237441	-0,721449

The confidence interval for the effect of sex we get by JMP is [-1,2374;-0,7214]. This parameter measures the difference in shoe size for a female and a male of the same height and the 95% confidence interval shows that we are 95% certain that the value of the effect of sex is within this interval. This also means that we are 95% that females have smaller shoe sizes than males.