Deskriptive Statistics (Gruppe2)

* Recruited subjects of the cohort: n=1201
* **Age by gender**
* **Puberty grouping beschreiben? (Körner S2)**
* **Puberty groups by gender**
* **Puberty groups by age & gender**
* Menarchal age , für Jungen?
* **BMI/BMI (SDS) by age**
* Height (SDS) by age & gender
* Weight (SDS) 🡪nur in Tabelle
* **Social status**
* **LH & FSH by age and gender**
* **Parameter (mean-) values by cohort**
* **LH/FSH auf Normalverteilung\* prüfen** (kolgorov-test or Shapiro-Wilk test )
* **LH/FSH Boxplots (nach kohorten, alter, PG, sozStat?) nach Kohorten nebeneinander** (Patricia)--> gucken ob statistisch signifikante Unterschiede zwischen SEX🡪 Mann-Whitney-Wilcoxon test/t-test( for Continuous), chi-quadrad für numerisch
* **Dependencies between age and parameter values**
* Am Ende: Übersichtstabelle mit allen means (Bußler), allen frequencies (elmlinger) machen

\*Statistisches Vorgehen

* Konforte et al: Prüfen auf Normalverteilung🡪 if not skewed: remove outliers using tukey test, if skewed: outliers identified and removed with an adjusted Tukey test.
* Soldin et al.: nutzt Hoffmann-Methode (?)🡪 von jeder Alterskathegorie werden outliers abgezogen

“The data were analyzed employing a computer adapted Hoffmann approach [[9]](http://www.sciencedirect.com/science/article/pii/S0009898105000501?via%3Dihub" \l "bib9). The data sets were separated into female and male subjects and stratified by age. Abnormal and outlier values were truncated from each individual age category according to the Hoffmann method. Generally, the top and bottom 10–20% of the data were discarded and the central linear portion of the graph extrapolated. The remaining data were either of normal Gaussian distribution or made to have a Gaussian distribution by calculating the logarithm of the values to determine the 2.5th and 97.5th percentiles for each of the age groups. Percent cumulative frequency versus concentration was plotted to calculate the 2.5th and 97.5th percentiles. These were used as the final reported serum concentration intervals.”

* Elmlinger et al.: “Age- or sex-specific differences between the groups were analyzed by means of the Mann and Whitney U-test. The detection limit was defined as the concentration two standard deviations above the response at zero dose.”
* Lindhardt-Johanson (Juul):
* The Mann-Whitney U test was used to compare the age at initial B2; the time of progression from B2 to B4; height velocity; height; weight; body mass index prepubertally, at pubertal onset, and postpubertally; and reproductive hormone levels in girls with and without TT.
* Wilcoxon’s signed-rank test was used to compare LH, FSH, estradiol, testosterone, … at the initial breast budding compared with the second breast budding
* Jensen et al. (Juul): LH in bezug auf PG🡪 unterschiede zwischen den Kohorten mit Kruskal wallis test.
* Mendiola et al. (Reproductive parameters in young men living in Rochester, New York

„Regarding reproductive hormones, the frequency distributions of serum hormone concentrations showed skewed (non-normal) distributions and were transformed using the natural log (ln) before analysis (estimates and 95% CIs were back-transformed as well). Covariates initially taken into account were age, BMI, ethnicity, smoking, season of sample collection, time of sample collection (number of hours after 12:00 PM of serum collection), and number of stressful life events. All tests were two-tailed, and the level of statistical significance was set at .05. Once final models were determined, two analysts (J.M. and F.L.) conducted these analyses independently using SAS version 8 (SAS Institute) and SPSS version 20.0 (IBM SPSS Statistics).“

* Recruited subjects of the cohort: n=1201

|  |  |
| --- | --- |
| female  male | 632  569 |
| sum | 1201 |

* **Age by gender**

|  |  |  |
| --- | --- | --- |
| Age | 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 | Sum |
| female  male | 0 0 49 65 41 54 67 55 63 65 49 47 34 28 15  1 0 36 55 42 71 69 60 59 57 44 36 22 9 8 | 632  569 |
| all | 1 0 85 120 83 125 136 115 122 122 93 83 56 37 23 | 1201 |

Table: Frequencies of age by gender

|  |  |
| --- | --- |
| female  male | n mean sd median trimmed mad min max range skew kurtosis se  632 10.78 3.28 10.68 10.69 3.9 5.54 17.95 12.41 0.16 -0.97 0.13   1. 10.38 2.97 10.25 10.28 3.28 3.21 17.88 14.67 0.25 -0.69 0.12 |

Table: Descriptive statistics of age by gender (1)

* Altersverteilung: Relativ normalverteilt (skew: leicht schief nach

links, kurtosis: fast bei 0 also relativ normal, leicht abgeflacht.

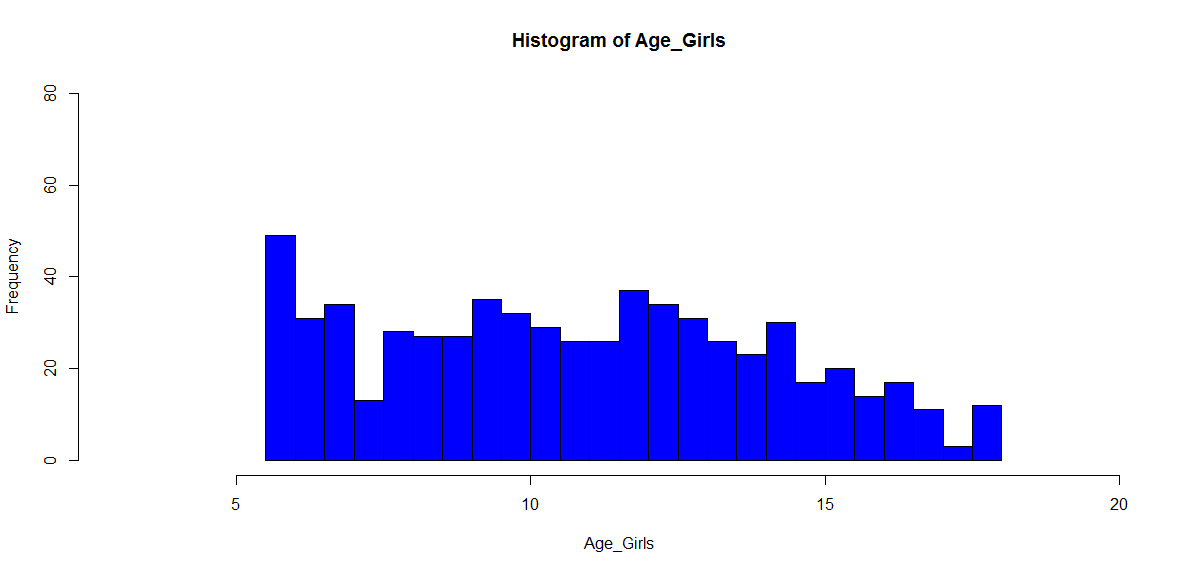
Teilt man durch die SD, ist der Wert nicht >1,96, also normal). Der

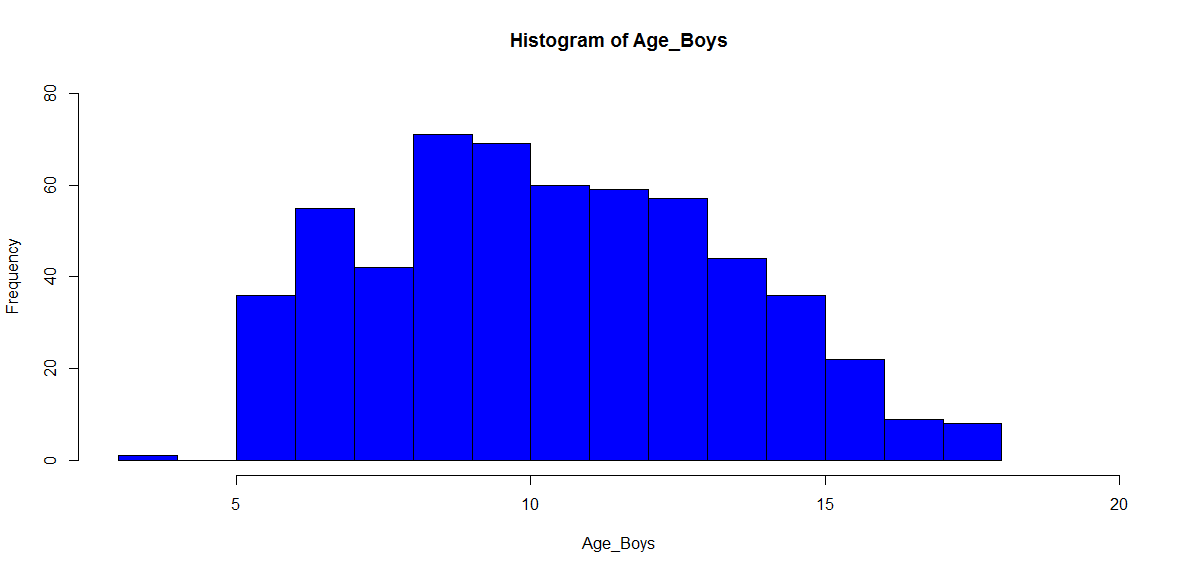
SE streut leicht um den Mittelwert, die SD zeigt eine große Variabilität

Der Daten an. Die Mädchen sind durschnittlich 10,78 Jahre alt, die Jungen 10,38 Jahre alt. *🡪 richtig so? Alle angaben beibehalten?*

|  |  |
| --- | --- |
| female  male | Min. 1st Qu. Median Mean 3rd Qu. Max.  5.540 8.072 10.680 10.783 13.312 17.950  3.21 8.11 10.25 10.38 12.53 17.88 |

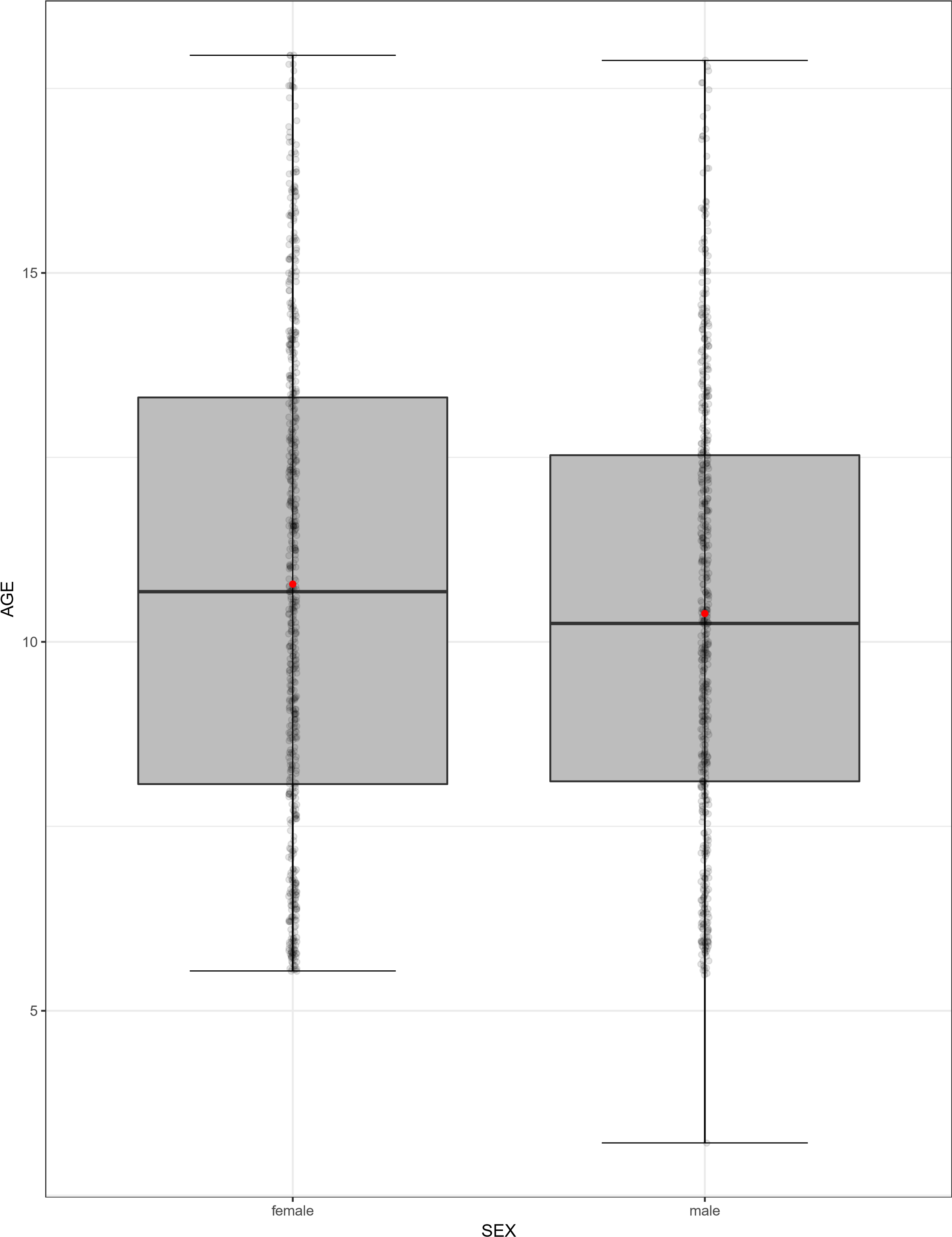
Table: Descriptive statistics of age by gender (2)





* *Ändern: nach einzelnen Jahren anzeigen, Titel ändern*

|  |  |
| --- | --- |
|  | |
|  |  |

*fragen: wie Achsenbezeichnungen größer?*

* ***Puberty grouping beschreiben? (Körner S2)***
* **Puberty groups by gender**

|  |  |  |
| --- | --- | --- |
|  | PG1 PG2 PG3 PG4 PG5 | Sum |
| female | 266 101 87 90 88 | 632 |
| male | 325 111 40 49 44 | 569 |
| all | 591 212 127 139 132 | 1201 |

Table: Frequencies of puberty groups by gender.

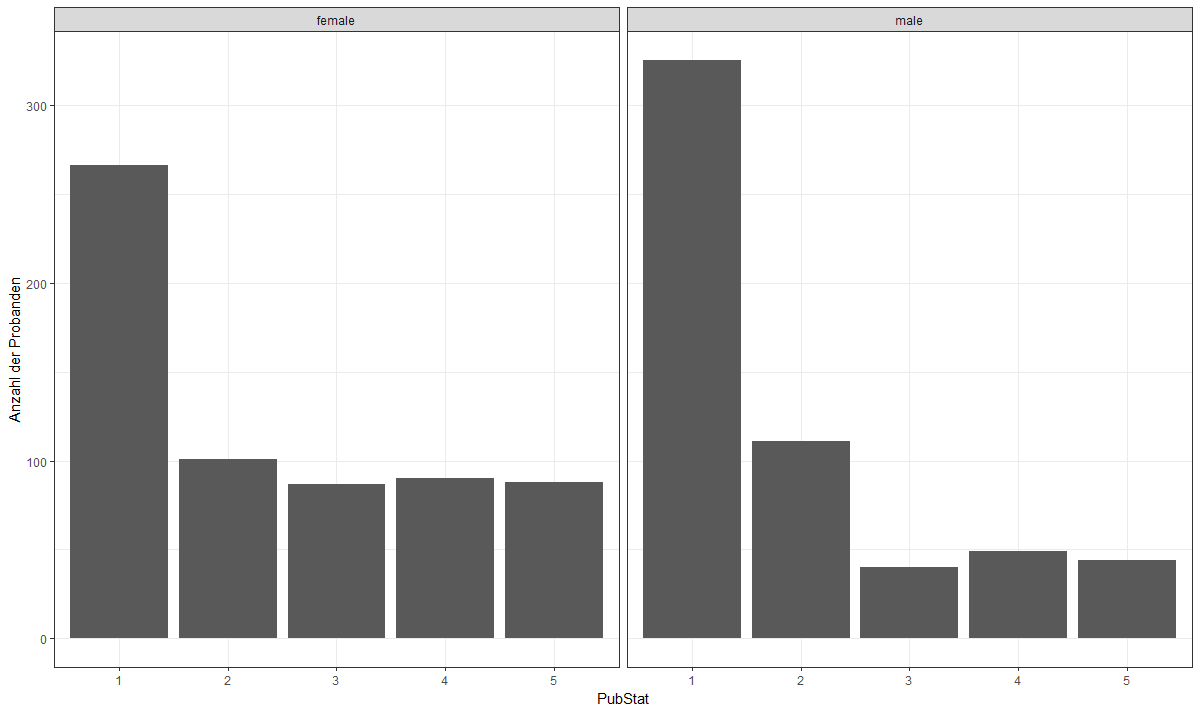


Figure: Graphs of puberty groups by gender

* **Puberty groups by age & gender**

|  |
| --- |
| age 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 Sum |
| PG1 1 0 36 55 42 71 52 35 19 12 2 0 0 0 0  PG2 0 0 0 0 0 0 16 23 30 29 11 2 0 0 0  PG3 0 0 0 0 0 0 1 2 7 10 11 9 0 0 0  PG4 0 0 0 0 0 0 0 0 3 6 16 16 6 2 0  PG5 0 0 0 0 0 0 0 0 0 0 4 9 16 7 8 |

Table: Frequencies of males‘ age by puberty groups

|  |
| --- |
| age 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 Sum |
| PG1 0 0 49 65 41 49 33 17 9 3 0 0 0 0 0  PG2 0 0 0 0 0 4 30 28 24 12 1 2 0 0 0  PG3 0 0 0 0 0 1 4 10 24 24 15 6 2 1 0  PG4 0 0 0 0 0 0 0 0 5 22 17 22 14 8 2  PG5 0 0 0 0 0 0 0 0 1 4 16 17 18 19 13 |

Table: Frequencies of females‘ age by puberty groups

* *Deskriptive Statistik- wie?? S9.*
* *wie kann ich noch rechts: sum und links unten „all females“ einfügen? (Körner S.8)*
* *wie kann ich eine Tabelle nicht manuell erstellen?*

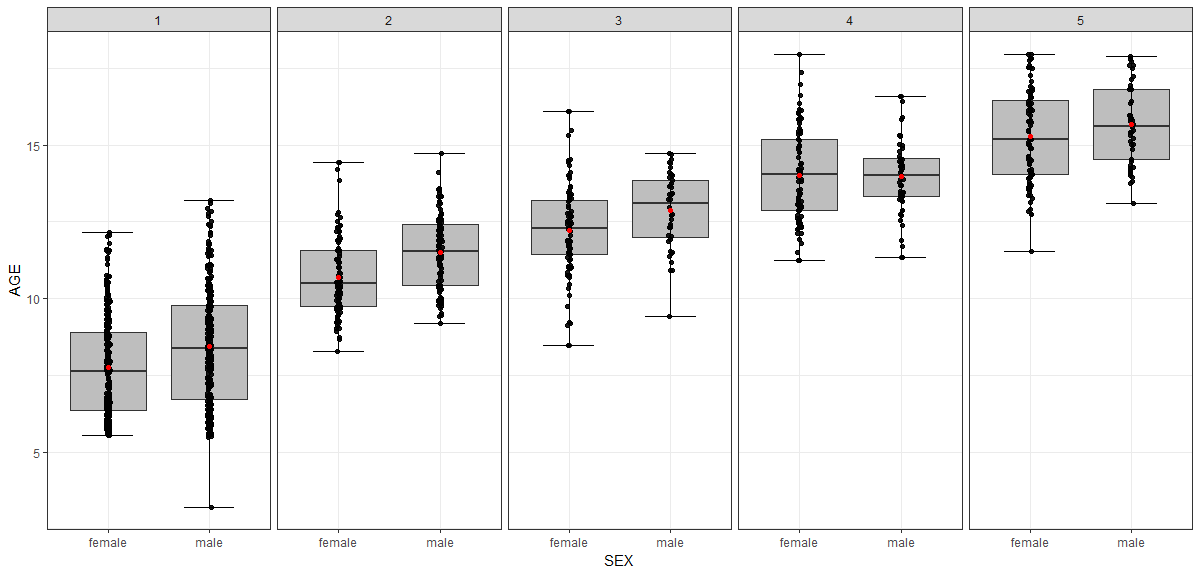


Figure: Boxplots of age by puberty groups and gender

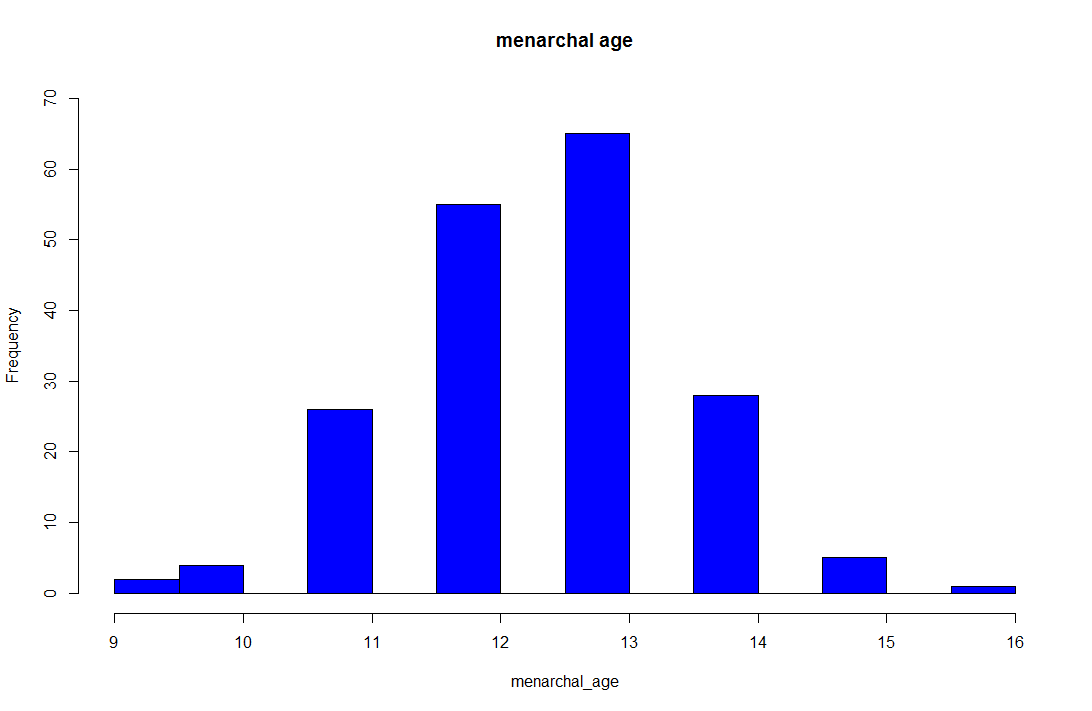
* **Menarchal age** (*und für Jungen?)*

|  |  |
| --- | --- |
| Menarchal age | 9 10 11 12 13 14 15 16 Sum |
| frequency | 2 4 26 55 65 28 5 1 186 |

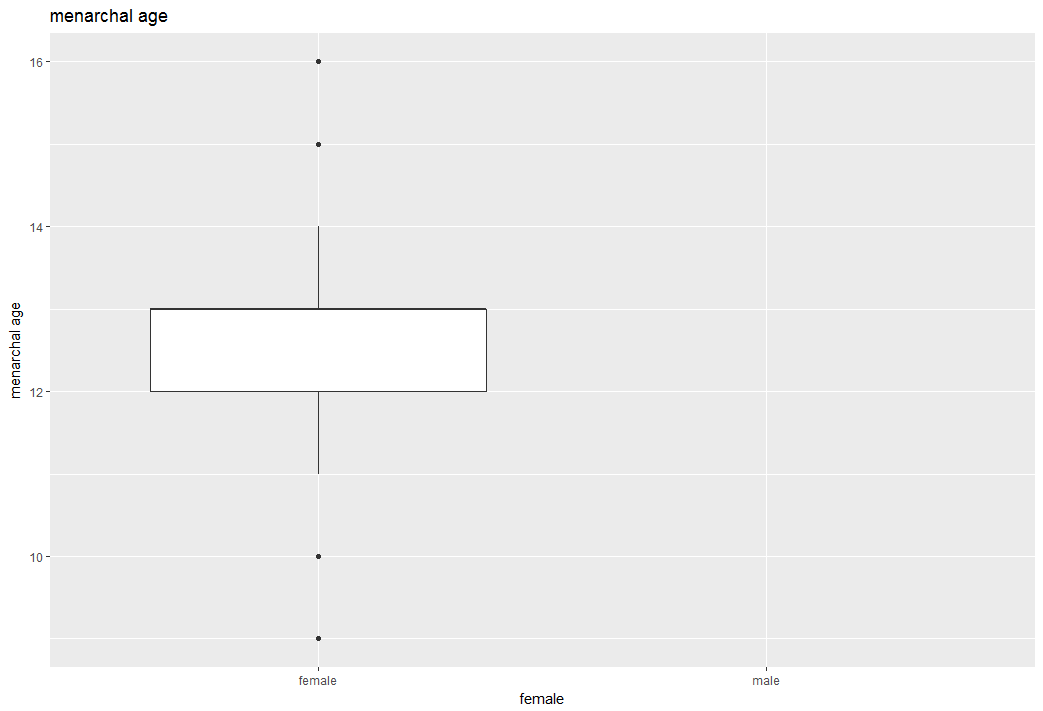
Table: frequency by menarchal age

|  |
| --- |
| n mean sd median trimmed mad min max range skew kurtosis se  186 12.54 1.15 13 12.55 1.48 9 16 7 -0.17 0.35 0.08 |

Table: Descripive statistics of menarchal age



* *Achsenbeschriftung ändern?*



* **BMI/BMI (SDS) by age & sex**

|  |  |
| --- | --- |
| female  male | n mean sd median trimmed mad min max range skew kurtosis se  628 19.12 4.68 17.79 18.51 3.76 11.78 41.86 30.08 1.42 2.63 0.19  568 18.9 4.82 17.36 18.12 3.16 12.93 53.05 40.12 1.83 5.2 0.2 |

Table: Descripitve statistics(1) of BMI by gender

|  |  |
| --- | --- |
| female  male | Min. 1st Qu. Median Mean 3rd Qu. Max. missing values N  11.78 15.69 17.79 19.12 21.38 41.86 4 628  12.93 15.48 17.36 18.90 20.68 53.05 1 568 |

Table: Descripitve statistics (2) of BMI by gender

|  |  |
| --- | --- |
| female  male | n mean sd median trimmed mad min max range skew kurtosis se  628 0.25 1.24 0.11 0.2 1.06 -4.62 3.52 8.15 0.24 0.27 0.05  568 0.26 1.19 0.04 0.2 1.11 -2.71 3.93 6.64 0.46 -0.21 0.05 |

Table: Descripitve statistics (1) of BMI (SDS) by gender

|  |  |
| --- | --- |
| female  male | Min. 1st Qu. Median Mean 3rd Qu. Max. missing values N  -4.6210 -0.5765 0.1110 0.2462 0.9842 3.5250 4 628  -2.7140 -0.5813 0.0405 0.2603 0.9645 3.9260 1 568 |

Table: Descripitve statistics (2) of BMI (SDS) by gender

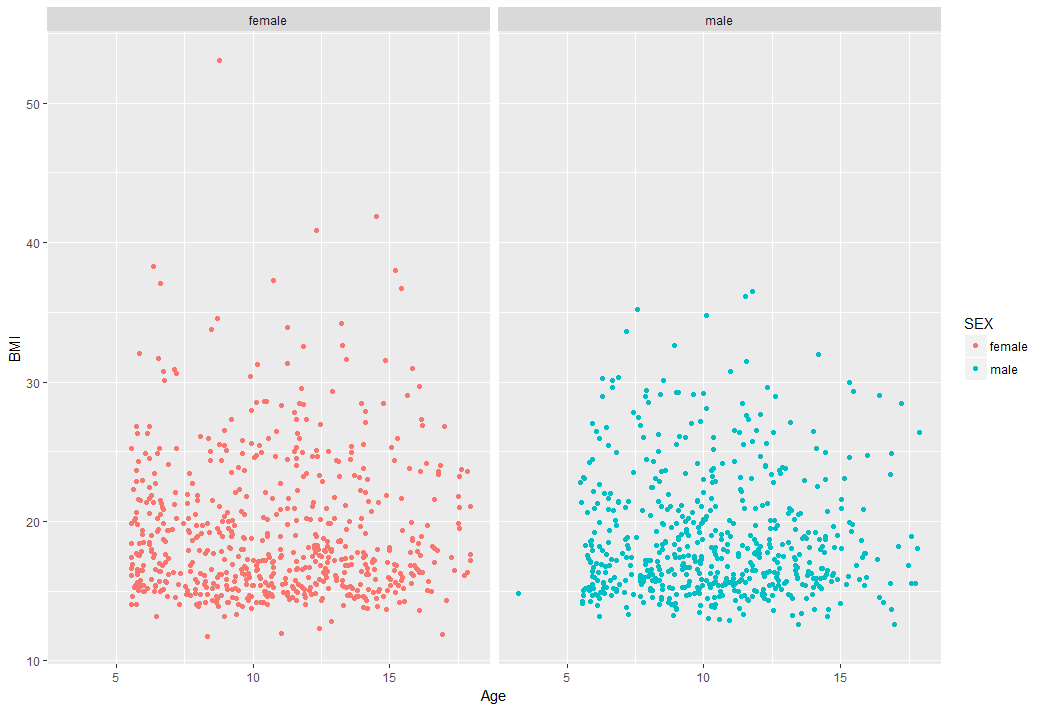


Figure: Scatterplots of age by BMI

*(skalierung: wie alter in 2 jahres stufen anzeigen lassen? Text (N=xx) einfügen rechts unten? Wie nach geschlecht getrennt einzelne grafen?) outliers?evtl. y-Achse limitieren?*

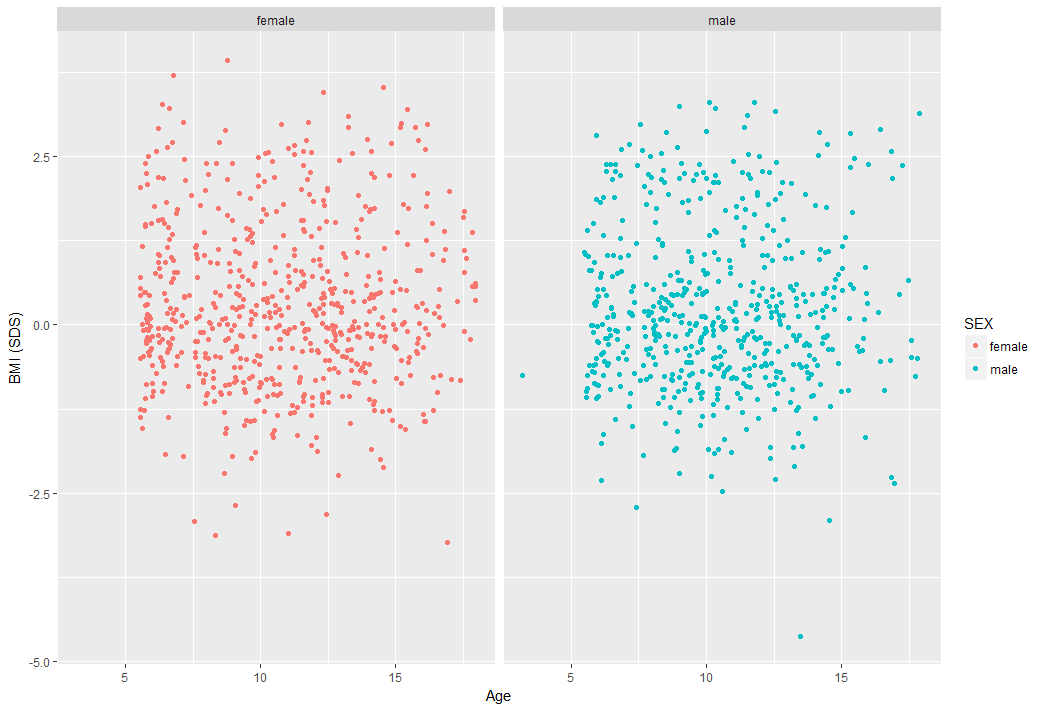


Figure: Scatterplots of age by BMI (SDS) 🡪*wichtig?*

* BMI: jetzt in normal, unter-, übergewichtig, adipös einteilen
* BMI distribution in Tabelle (siehe Bußler)
* **Height by age & gender** *--> auch SDS?*

|  |  |
| --- | --- |
| female  male | n mean sd median trimmed mad min max range skew kurtosis se  629 145.46 18.33 147.3 146.08 22.54 104 182.7 78.7 -0.24 -1.08 0.73  568 145.52 19.33 144.8 145.01 19.87 95.1 196.3 101.2 0.2 -0.66 0.81 |

Table: Descripitve statistics (1) of height by gender

|  |  |
| --- | --- |
| female  male | Min. 1st Qu. Median Mean 3rd Qu. Max. missing values N  11.78 15.69 17.79 19.12 21.38 41.86 4 629  12.93 15.48 17.36 18.90 20.68 53.05 1 568 |

Table: Descripitve statistics (2) of height by gender

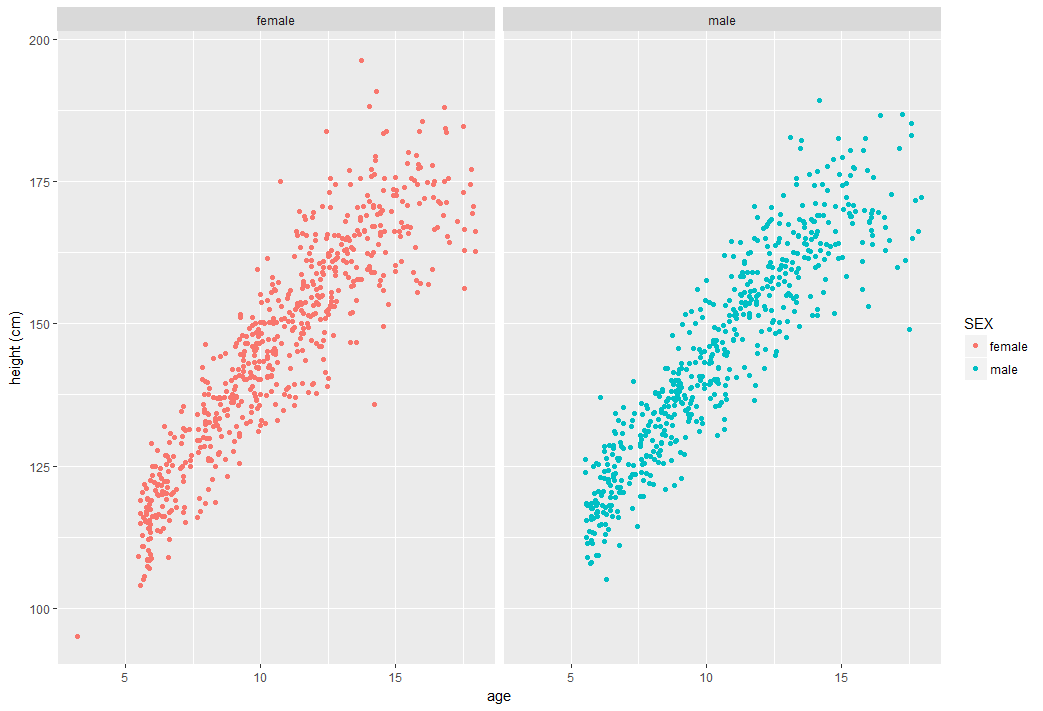
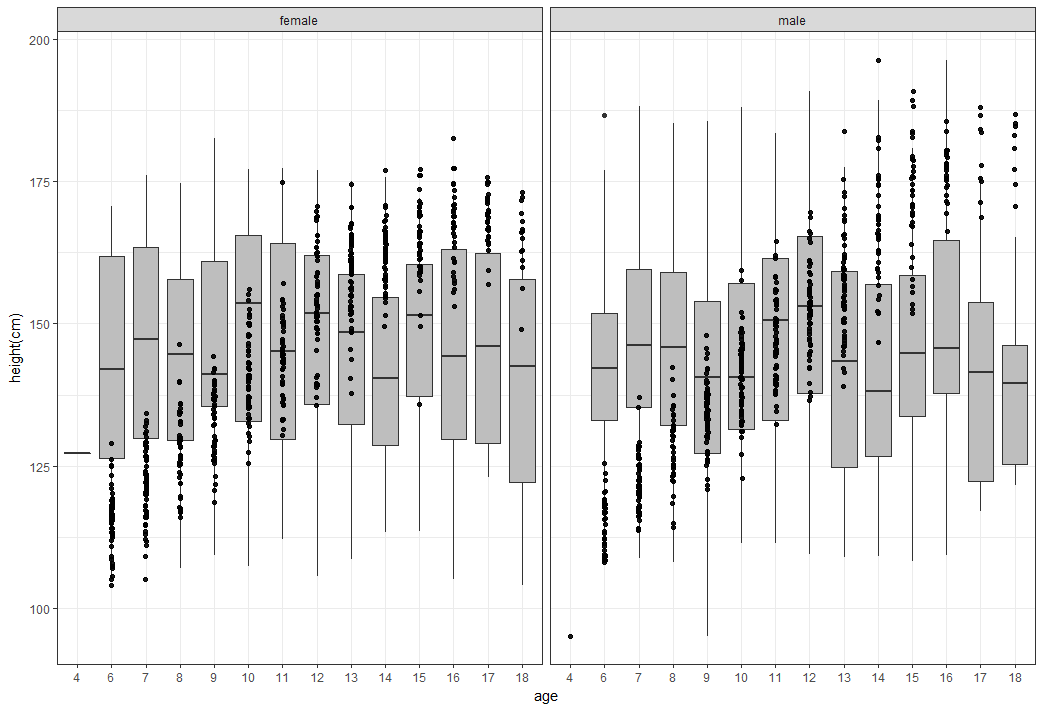


Figure: Scatterplots of age by height

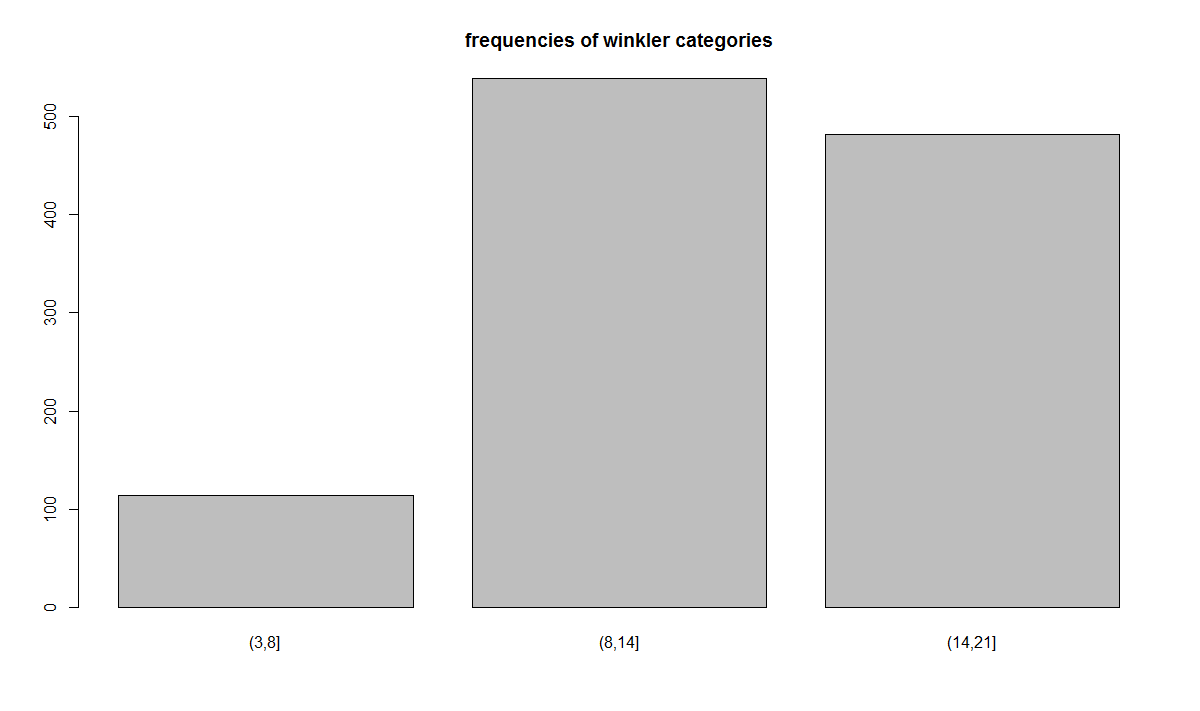
🡪*Achsenskalierung: 6,8,10,etc.? Text hinzufügen in Plots (n=xx)? noch SDS? Achsenbeschriftung größer?*

*?? Macht wohl eher keinen Sinn…*

* ***Kohorten*** *nahc BMI einteilen! In normal/übergewichtig/adipös*
* *Mean von Alter, BMI, SDS in boxplots nebeneinander stellen*
* **Social status**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Social status | low | intermediate | high | missing values | sum |
| Anzahl | 114 | 538 | 481 | 68 | 1201 |

* *Bezeichnung social status?*



* **LH & FSH by gender**
* LH

|  |  |
| --- | --- |
| female  male | n mean sd median trimmed mad min max range skew kurtosis se  628 2.82 5.35 0.11 1.63 0.01 0.1 46.61 46.51 3.82 19.23 0.21  562 1.05 1.5 0.1 0.73 0 0.1 7.81 7.71 1.82 3.07 0.06 |

Table: Descripitve statistics (1) of LH by gender

|  |  |
| --- | --- |
| female  male | Min. 1st Qu. Median Mean 3rd Qu. Max. missing values N  0.100 0.100 0.105 2.823 3.875 46.610 4 628  0.100 0.100 0.100 1.051 1.657 7.810 7 562 |

Table: Descripitve statistics (2) of LH by gender

* *Warum male n=562*
* FSH

|  |  |
| --- | --- |
| female  male | n mean sd median trimmed mad min max range skew kurtosis se  632 3.38 2.31 2.63 3.11 2.05 0.14 12.15 12.01 1 0.44 0.09  569 1.89 1.69 1.35 1.61 1.1 0.2 12.75 12.55 2.19 7.19 0.07 |

Table: Descripitve statistics (1) of FSH by gender

|  |  |
| --- | --- |
| female  male | Min. 1st Qu. Median Mean 3rd Qu. Max. N  0.140 1.518 2.630 3.380 4.952 12.150 632  0.200 0.700 1.350 1.895 2.530 12.750 569 |

Table: Descripitve statistics (2) of FSH by gender

* **Parameter (mean-) values by cohort**
* *Andere parameter mergen (SHBG, Estradiol, Testosteron, DHEAS)*
* **Dependencies between age and parameter values**

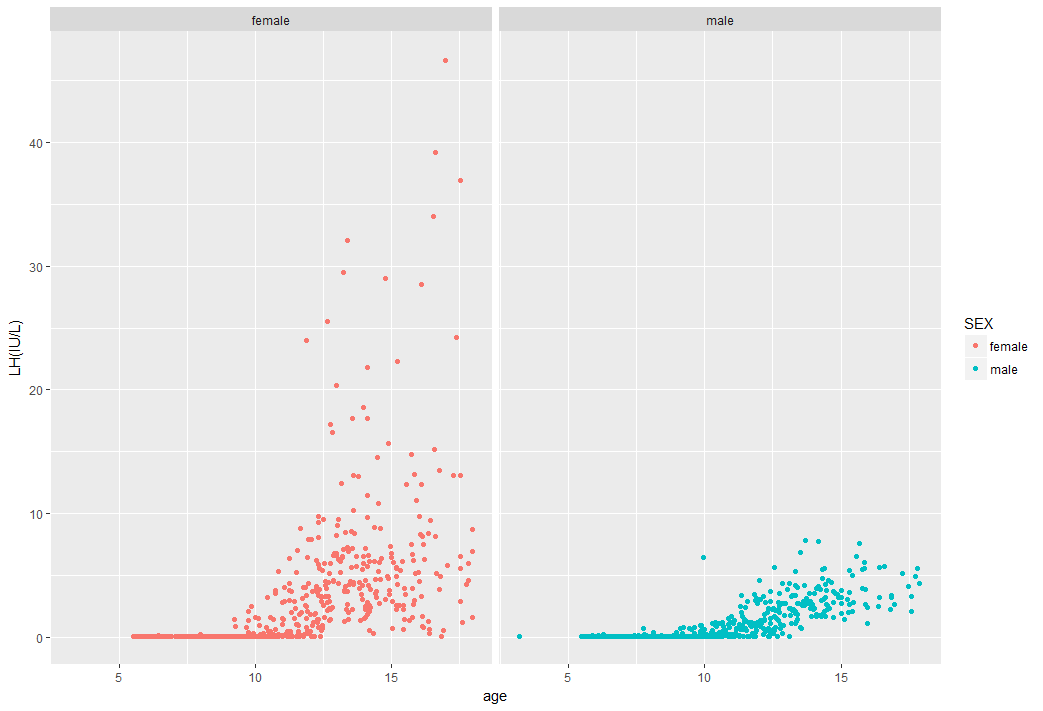


Figure: Scatterplots of age by LH

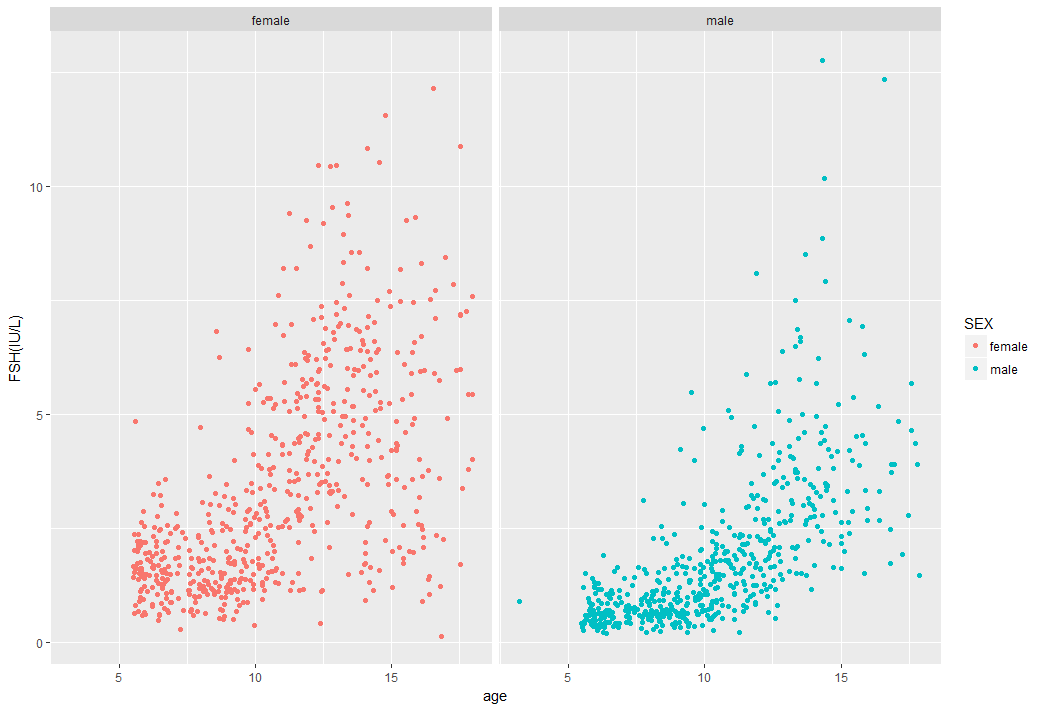


Figure: Scatterplots of age by FSH

* LH/FSH by puberty groups

Fragen Thomas:

* Puberty grouping beschreiben? Siehe Körner S.2. Wie haben bußler und co es gemacht? Bußler hat in 2 Gruppen eingeteilt (Tanner1-2, 3-5).
* Menarchal age- und für Jungen? 🡪 juul nachgucken, …
* Alles auch SDS?!--> werte einspielen
* Nur wenige übergewichtige…weitere Daten? Wie soll ich es machen?
* Sachen durchgehen (word-plan, dokument und Fragen, und R)

Patricia

* Einteilen: overweight etc: underweight ausschließen?? Gucken wieviele.
* Test: LH/FSH normalverteilung: kolgorov-test or Shapiro-Wilk test. Und Graph dazu. (density~lh)
* Boxplot: LH für alle kohorten nebeneinander und vergleichen: tests je nachdem ob normalverteilt: normal=t-test, nicht-normal: mann whitney-test (für continuous, numerical), für kategorisch: chi-quadrad
* An anderen Publikationen orientieren
* ANOVA: for more than 2, always+post-hoc test. otherwise: students t-test,