How does a student’s data analysis process change over time?

A block-level analysis

## The Block Model (Schulte 2008)

|  | Text Surface | Program Execution | Function |
| --- | --- | --- | --- |
| **Macrostructure** | Understanding the overall structure of the program | Understanding the “algorithm” of the program | Understanding the goal / purpose of the program (in its context) |
| **Relations** | References between blocks, e.g., method calls, object creation | Sequence of method calls, object sequence diagrams | Understanding how sub-goals are related to goals, how function is achieved by subfunctions |
| **Blocks** | Regions of interest (ROI) that syntactically or semantically build a unit | Operation of a block, a method, or a ROI (as a sequence of statements) | Function of a block, may be seen as a sub-goal |
| **Atoms** | Language elements | Operation of a statement | Function of a statement, only understandable in context |

## Time Point 1

#upper anterior measurement  
anterior <- lm(ProximateAnalysisData$PSUA~ProximateAnalysisData$Lipid)  
summary(anterior)  
with(ProximateAnalysisData, plot(PSUA ~ Lipid, las = 1))  
abline(anterior)  
plot(anterior)  
  
posterior <- lm(ProximateAnalysisData$PSUP ~ ProximateAnalysisData$Lipid)  
summary(posterior)  
posterior  
with(ProximateAnalysisData, plot(PSUP ~ Lipid, las = 1))  
abline(posterior)  
plot(posterior)  
  
#OUTLIER REMOVED  
anterior2 <- lm(ProximateAnalysisDataOutlier$PSUA ~ ProximateAnalysisDataOutlier$Lipid)  
summary(anterior2)  
with(ProximateAnalysisDataOutlier, plot(PSUA ~ Lipid, las = 1, xlab = "Whole-body Lipid Content (%)", ylab = "UA Fatmeter Reading"))  
abline(anterior2)  
plot(anterior2)  
anterior2  
  
posterior2 <- lm(ProximateAnalysisDataOutlier$PSUP ~ ProximateAnalysisDataOutlier$Lipid)  
summary(posterior2)  
with(ProximateAnalysisDataOutlier, plot(PSUP ~ Lipid, las = 1, xlab = "Whole-body Lipid Content (%)", ylab = "UP Fatmeter Reading"))  
abline(posterior2)  
plot(posterior2)  
posterior2  
  
qt(.975,9)  
  
#upper Middle measurements only  
middle <- lm(ProximateAnalysisData$PSUM ~ ProximateAnalysisData$Lipid)  
summary(middle)  
with(ProximateAnalysisData, plot(PSUM ~ Lipid, las = 1, xlab = "Whole-body Lipid Content (%)", ylab = "UM Fatmeter Reading"))  
abline(middle)  
plot(middle)  
middle  
  
middleLog <- lm(ProximateAnalysisData$logPSUM ~ ProximateAnalysisData$Lipid)  
summary(middleLog)  
with(ProximateAnalysisData, plot(logPSUM ~ Lipid, las = 1))  
abline(middleLog)  
plot(middle)  
with(ProximateAnalysisData, plot(Lipid ~ logPSUM, las = 1))  
  
#Means and sd of data  
mean(ProximateAnalysisData$Lipid)  
sd(ProximateAnalysisData$Lipid)

## Time Point 2

#Preliminary data for Fatmeter Calibration  
#-----------------------------------------------  
  
#-----------------------------  
# WB Lipid Analysis  
#-----------------------------  
  
#upper anterior measurement(outlier included) Linear model  
#Only fish that had upper anterior measurements were the larger cohort sampled at Garrison  
  
linearAnterior <- lm(ProximateAnalysisData$PSUA~ProximateAnalysisData$Lipid)  
summary(linearAnterior)  
linearAnterior  
with(ProximateAnalysisData, plot(PSUA ~ Lipid, las = 1))  
abline(linearAnterior)  
plot(linearAnterior)  
  
#Exponential function  
expAnterior <- lm(log(ProximateAnalysisData$PSUA)~ProximateAnalysisData$Lipid)  
summary(expAnterior)  
expAnterior  
with(ProximateAnalysisData, plot(log(PSUA) ~ Lipid, las = 1))  
  
#Upper posterior measurement (outlier included)  
#Only fish that had upper posterior measurements were the larger cohort sampled at Garrison  
  
linearposterior <- lm(ProximateAnalysisData$PSUP ~ ProximateAnalysisData$Lipid)  
summary(linearposterior)  
linearposterior  
with(ProximateAnalysisData, plot(PSUP ~ Lipid, las = 1))  
plot(posterior)  
  
#Exponential posterior measurement (outlier included)  
expPosterior <- lm(log(ProximateAnalysisData$PSUP) ~ ProximateAnalysisData$Lipid)  
summary(expPosterior)  
expPosterior  
with(ProximateAnalysisData, plot(log(PSUP) ~ Lipid), las = 1)  
plot(expPosterior)  
  
#Anterior measuremnet with OUTLIER REMOVED exponential  
expAnterior2 <- lm(log(ProximateAnalysisDataOutlier$PSUA) ~ ProximateAnalysisDataOutlier$Lipid)  
summary(expAnterior2)  
expAnterior2  
with(ProximateAnalysisDataOutlier, plot(log(PSUA) ~ Lipid), las = 1, xlab = "Whole-body Lipid Content (%)", ylab = "UA Fatmeter Reading")  
abline(expAnterior2)  
plot(anterior2)  
anterior2  
  
#Posterior measuremnet with OUTLIER REMOVED  
expPosterior2 <- lm(log(ProximateAnalysisDataOutlier$PSUP) ~ ProximateAnalysisDataOutlier$Lipid)  
summary(expPosterior2)  
with(ProximateAnalysisDataOutlier, plot(log(PSUP) ~ Lipid, las = 1))  
abline(expPosterior2)  
plot(posterior2)  
posterior2  
  
#CI  
qt(.975,9)  
  
#upper Middle measurements only (Including outlier)  
expMiddle <- lm(ProximateAnalysisData$PSUM ~ log(ProximateAnalysisData$Lipid))  
summary(expMiddle)  
with(ProximateAnalysisData, plot(PSUM ~ log(Lipid), las = 1, xlab = "Whole-body Lipid Content (%)", ylab = "UM Fatmeter Reading"))  
plot(middle)  
middle  
  
#---------------------------------  
#Energy analysis of data  
#---------------------------------  
  
#Means and sd of data  
mean(ProximateAnalysisData$Energy)  
sd(ProximateAnalysisData$Energy)  
  
#Anterior energy measurement (outlier included)  
expAnteriorE <- lm(log(ProximateAnalysisData$PSUA)~ProximateAnalysisData$Energy)  
summary(expAnteriorE)  
with(ProximateAnalysisData, plot(log(PSUA) ~ Energy, las = 1))  
plot(anteriorE)  
  
#Posterior energy measurement (outlier included)  
expPosteriorE <- lm(log(ProximateAnalysisData$PSUP) ~ ProximateAnalysisData$Energy)  
summary(expPosteriorE)  
expPosteriorE  
with(ProximateAnalysisData, plot(PSUP ~ Energy, las = 1))  
plot(posteriorE)  
  
#OUTLIER REMOVED anterior Energy  
expAnterior2E <- lm(log(ProximateAnalysisDataOutlier$PSUA) ~ ProximateAnalysisDataOutlier$Energy)  
summary(expAnterior2E)  
expAnterior2E  
with(ProximateAnalysisDataOutlier, plot(log(PSUA) ~ Energy, las = 1))  
abline(expAnterior2E)  
plot(anterior2E)  
  
#Posterior energy (outlier removed)  
expPosterior2E <- lm(log(ProximateAnalysisDataOutlier$PSUP) ~ ProximateAnalysisDataOutlier$Energy)  
summary(expPosterior2E)  
with(ProximateAnalysisDataOutlier, plot(log(PSUP) ~ Energy, las = 1))  
abline(expPosterior2E)  
plot(posterior2E)  
posterior2E  
  
#CI  
qt(.975,9)  
  
#Middle Data (outlier removed)  
expMiddle2E <- lm(log(ProximateAnalysisDataOutlier$PSUM) ~ ProximateAnalysisDataOutlier$Energy)  
summary(expMiddle2E)  
expMiddle2E  
with(ProximateAnalysisDataOutlier, plot(PSUM ~ log(Energy), las = 1))  
plot(middle2E)

Schulte, Carsten. 2008. “Block Model.” *Proceedings of the Fourth International Workshop on Computing Education Research*, September. <https://doi.org/10.1145/1404520.1404535>.