|  |
| --- |
| Final Cleaning |
| Starting files:  Data1415: Year1.csv  Data1516: Year2.csv  Data1617: Year3.csv |
| Renaming and recoding variables  Remove opt out  Separate into condensed and regular |
| Ending files:  All Years Final v3.csv  All Years Final v3 Public.csv (codes in Institution\_Name\_Anonymized)  Condensed Final v3.csv |
| <<recoding the outcomes of the concept questions to be more uniform>> |
| Ending files:  ~~All Years Final.csv~~  All Years Final Public.csv  Condensed Final.csv |

|  |
| --- |
| Optional Subsetting |
| All Years Final Public.csv |
| Optin (to both)?  Remove low response rate (80%)  Remove extreme achieve gain (-1.1)  ~~Remove ISCAM and own~~  Remove calc prereq and precalc prereq  Remove HS (Carnegie)  Recoding  Section level variables |
| Post Optional Subsetting.csv |

|  |
| --- |
| Optional Imputation |
| Post Optional Subsetting.csv |
| Imputation (Amelia package) |
| Post Optional Subsetting Imputed.csv |

Questions

* out vs.out1? *consequence of amelia*
* Recode sooner?
* Pull out imputation?
  + Pull out section level coding?

|  |
| --- |
| Data Exploration Ajay  Data Exploration Aug 10 |
| Post Optional Subsetting.csv or Post Optional Subsetting Imputed.csv |
| Tables for paper  Boxplots  Question number table  Post Optional Subsetting Imputed.csv  StepAIC |
|  |

|  |
| --- |
| Data Analysis  Data Analysis Aug 31 |
| Post Optional Subsetting Imputed.csv |
| ModelBuilding |
|  |

Questions: file differences

|  |
| --- |
| AllVariables |
| Post Optional Subsetting Imputed  Post Optional Subsetting.csv |
| My model building? |
|  |