STEM Careers

DG

4/26/2021

Background

The purpose of this project is to “

Loading data

Load in the data sets for the

```
library(data.table)
library(dplyr)
library(ggplot2)
library(xlsx)
library(XML)
library(flextable)
library(tibble)
set_flextable_defaults(fonts_ignore=TRUE)

dataTotal <- data.frame(read.xlsx2("./DataInTotal.xlsx",
                                   colClasses=rep("numeric",24), sheetIndex = 1))
dataGender <- data.frame(read.xlsx2("./DataByGender.xlsx",
                                    colClasses=rep("numeric",25), sheetIndex = 1))

# change all zero and NaN values to NA in both datasets
dataTotal <- na_if(dataTotal, 0)
dataGender <- na_if(dataGender, 0)
dataTotal <- replace(dataTotal, is.na(dataTotal), NA)
dataGender <- replace(dataGender, is.na(dataGender), NA)
```

Summary Statistics for Full Data Sets

```
statsGender <- data.table()

Mean <- dataTotal %>% summarize(across(ArtsMatch:STEMMatch, ~ mean(.x, na.rm=TRUE)))
Median <- dataTotal %>% summarize(across(ArtsMatch:STEMMatch, ~ median(.x, na.rm=TRUE)))
SD <- dataTotal %>% summarize(across(ArtsMatch:STEMMatch, ~ sd(.x, na.rm=TRUE)))
statsTotal <- as.data.frame(cbind(t(Mean[1,]), t(Median[1,]), t(SD[1,])))
names(statsTotal) <- c("Mean", "Median", "SD")

MeanMale <- dataGender %>% filter(Gender==1) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                  ~ mean(.x, na.rm=TRUE)))
MedianMale <- dataGender %>% filter(Gender==1) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                    ~ median(.x, na.rm=TRUE)))
```

```
SDMale <- dataGender %>% filter(Gender==1) %>% summarize(across(ArtsMatch:STEMMatch,
                                                               ~ sd(.x, na.rm=TRUE)))
statsGender <- as.data.frame(cbind(t(MeanMale[1,]), t(MedianMale[1,]), t(SDMale[1,])))
names(statsGender) <- c("Mean", "Median", "SD")
row.names(statsGender) <- paste(row.names(statsGender), " (Male)")

MeanFemale <- dataGender %>% filter(Gender==2) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                   ~ mean(.x, na.rm=TRUE)))
MedianFemale <- dataGender %>% filter(Gender==2) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                     ~ median(.x, na.rm=TRUE)))
SDFemale <- dataGender %>% filter(Gender==2) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                 ~ sd(.x, na.rm=TRUE)))
statsGender2 <- as.data.frame(cbind(t(MeanFemale[1,]), t(MedianFemale[1,]), t(SDFemale[1,])))
names(statsGender2) <- c("Mean", "Median", "SD")
row.names(statsGender2) <- paste(row.names(statsGender2), " (Female)")
statsGender <- rbind(statsGender, statsGender2)
```

Table 1: Summary statistics for the data without gender included

```
ft<- flextable(statsTotal %>% rownames_to_column("Name"))
autofit(ft)
```

Name	Mean	Median	SD
ArtsMatch	0.11645079	0.13131326	0.057468751
BusMatch	0.32014908	0.42707453	0.168164890
CompDirect	0.09634964	0.10270006	0.040577975
CompIndir	0.39606368	0.43255539	0.150910597
CompSTEM	0.43884156	0.48316977	0.163219334
EduMatch	0.37400566	0.44353694	0.257186491
EngrDirect	0.08189763	0.09231602	0.041682104
EngrIndir	0.19743836	0.25121215	0.109768681
EngrSTEM	0.31783746	0.39664503	0.198315463
HealthMatch	0.41314987	0.54186097	0.234984233
LifeSciDirect	0.01297409	0.01223531	0.005766597
LifeSciIndir	0.05086392	0.05488805	0.023764831
LifeSciSTEM	0.11741079	0.13064671	0.058012694
MathDirect	0.02776403	0.02365599	0.020580572
MathIndir	0.05295519	0.05240165	0.033363198
MathSTEM	0.21119321	0.24777416	0.106598773
PhySciDirect	0.05833211	0.05942289	0.030797171
PhySciIndir	0.07842263	0.08596192	0.040111303
PhySciSTEM	0.18150532	0.22930403	0.098788147
SocSciMatch	0.13231193	0.16236972	0.061786477
ANYMATCH	0.23053391	0.30859610	0.133850865

Name	Mean	Median	SD
STEMtoSTEM	0.24117469	0.31789437	0.149584697
STEMMatch	0.09864362	0.12197610	0.060921534

Table 2: Summary statistics for the data broken out by gender

```
ft<- flextable(statsGender %>% rownames_to_column("Name"))
autofit(ft)
```

Name	Mean	Median	SD
ArtsMatch (Male)	0.14680363	0.16493040	0.065031518
BusMatch (Male)	0.34850909	0.45708780	0.190697532
CompDirect (Male)	0.10479568	0.11189684	0.039699908
CompIndir (Male)	0.43950338	0.49817296	0.160416220
CompSTEM (Male)	0.48843710	0.54851098	0.173321736
EduMatch (Male)	0.32304264	0.36257786	0.220034343
EngrDirect (Male)	0.08720663	0.09970499	0.045666748
EngrIndir (Male)	0.20843108	0.26480348	0.117319153
EngrSTEM (Male)	0.33683222	0.41803264	0.208639661
HealthMatch (Male)	0.42430862	0.50836606	0.176275314
LifeSciDirect (Male)	0.01549296	0.01551543	0.007888987
LifeSciIndir (Male)	0.05176282	0.05235544	0.024863777
LifeSciSTEM (Male)	0.13580759	0.14406229	0.078373681
MathDirect (Male)	0.02956399	0.02168004	0.024444785
MathIndir (Male)	0.05588522	0.05204973	0.036949945
MathSTEM (Male)	0.25361992	0.28619847	0.130419447
PhySciDirect (Male)	0.06797877	0.06893524	0.029651038
PhySciIndir (Male)	0.09003708	0.10013142	0.042797348
PhySciSTEM (Male)	0.22104398	0.27571735	0.122487386
SocSciMatch (Male)	0.11068778	0.12636071	0.042530574
ANYMATCH (Male)	0.22220765	0.29613800	0.120481295
STEMtoSTEM (Male)	0.28332964	0.36681514	0.179406099
STEMMatch (Male)	0.11843836	0.13399154	0.075008851
ArtsMatch (Female)	0.10385586	0.11235326	0.054089459
BusMatch (Female)	0.28700778	0.38195068	0.165706538
CompDirect (Female)	0.08705431	0.08061894	0.045856474
CompIndir (Female)	0.32334815	0.33749864	0.112242620
CompSTEM (Female)	0.35454516	0.36861322	0.116642750

Name	Mean	Median	SD
EduMatch (Female)	0.39642523	0.47027674	0.265519648
EngrDirect (Female)	0.06958699	0.07014507	0.032685858
EngrIndir (Female)	0.17949950	0.17984715	0.076340708
EngrSTEM (Female)	0.30760680	0.31359497	0.136199823
HealthMatch (Female)	0.43563466	0.56477762	0.237942129
LifeSciDirect (Female)	0.01204345	0.01087075	0.005688901
LifeSciIndir (Female)	0.05720413	0.05778269	0.022988797
LifeSciSTEM (Female)	0.11722383	0.12827612	0.049507101
MathDirect (Female)	0.03386979	0.02970828	0.024662220
MathIndir (Female)	0.05783779	0.05526834	0.039631139
MathSTEM (Female)	0.17244513	0.18227607	0.093252014
PhySciDirect (Female)	0.04877278	0.04426968	0.040195966
PhySciIndir (Female)	0.07376922	0.07091430	0.041783231
PhySciSTEM (Female)	0.14184913	0.15336134	0.069411041
SocSciMatch (Female)	0.15016193	0.18135349	0.077084377
ANYMATCH (Female)	0.24036155	0.33241617	0.145027979
STEMtoSTEM (Female)	0.17012397	0.20643267	0.096826504
STEMMatch (Female)	0.07604761	0.08399290	0.033090718

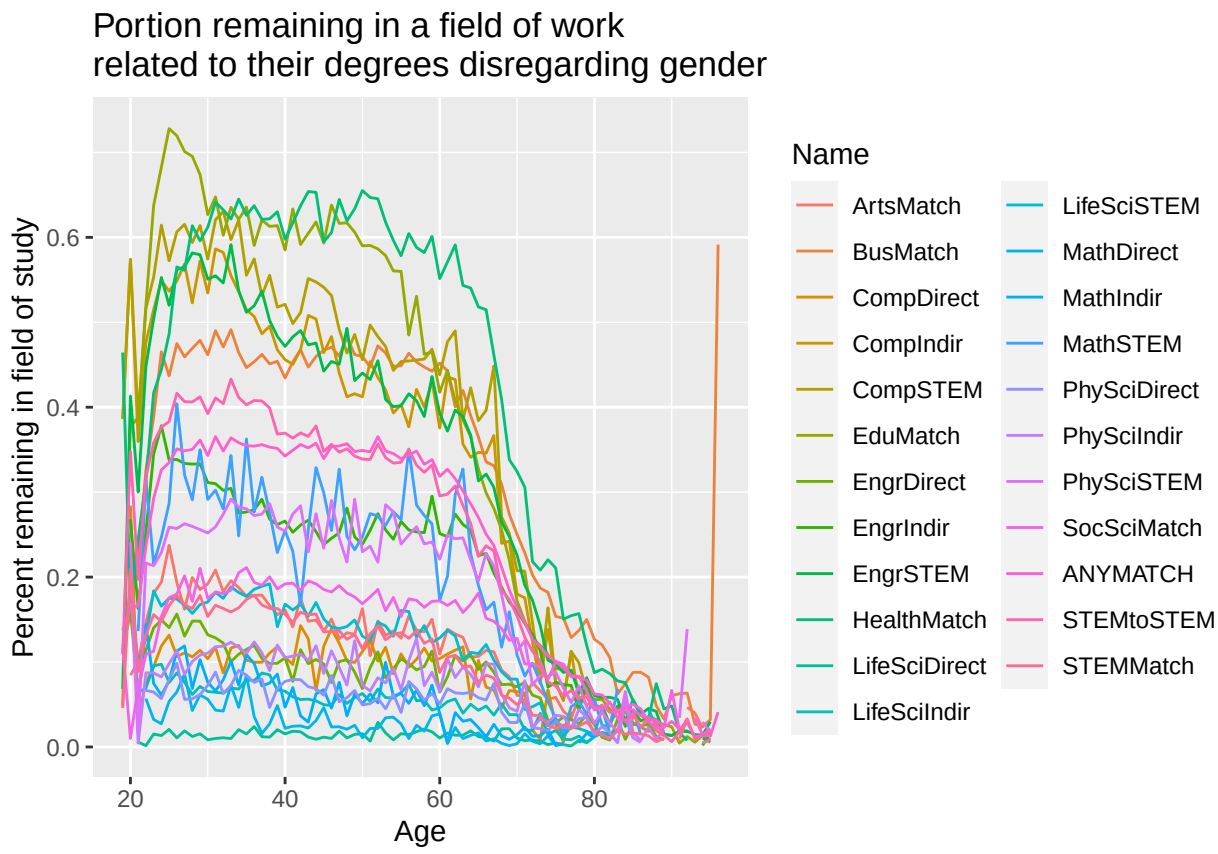
Graph All Degree-Types by Age

```
library(reshape2)
```

```
##
## Attaching package: 'reshape2'
```

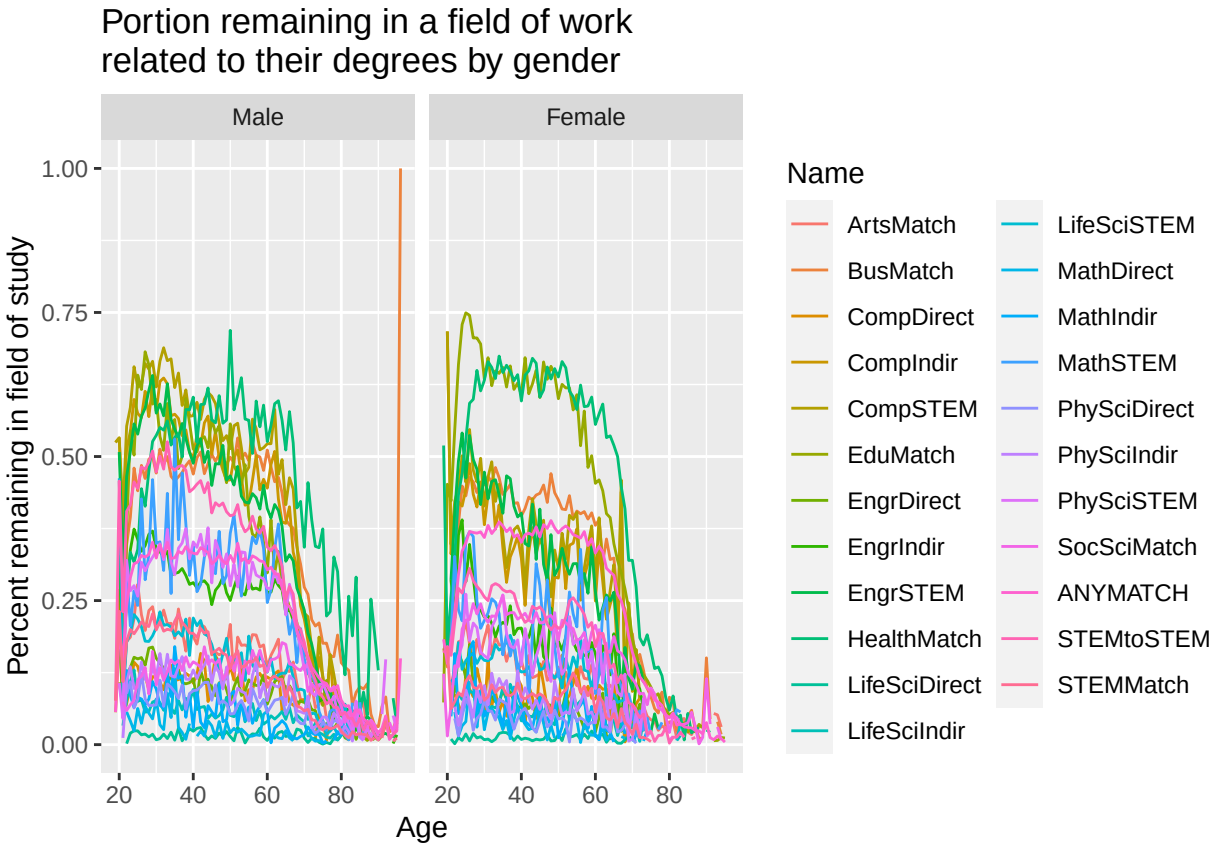
```
## The following objects are masked from 'package:data.table':
##
##     dcast, melt
```

```
library(ggplot2)
meltTotal <- reshape2::melt(dataTotal, id.vars="Age", variable.name="Name", value.name = "Percentage")
TotalPlot <- ggplot(meltTotal, aes(x=Age, y=Percentage, group=Name))
TotalPlot <- TotalPlot +  geom_line(aes(color=Name))
TotalPlot <- TotalPlot +  labs(title="Portion remaining in a field of work\nrelated to their degrees di
TotalPlot
```



Graph All Degree-Types by Age and by Gender

```
library(reshape2)
meltGender <- reshape2::melt(dataGender, id.vars=c("Age", "Gender"), variable.name="Name", value.name =
GenderPlot <- ggplot(meltGender, aes(x=Age, y=Percentage, group=Name))
GenderPlot <- GenderPlot +  geom_line(aes(color=Name))
GenderPlot <- GenderPlot +  labs(title="Portion remaining in a field of work\nrelated to their degrees
Gender.labs <- c("Male", "Female")
names(Gender.labs) <- c("1", "2")
GenderPlot <- GenderPlot +  facet_grid(. ~Gender, labeller = labeller(Gender = Gender.labs))
GenderPlot
```



Subsetting Data for Ages 30-65

The summary data indicated a high variance for some of the degree types and a wide range of variances across the degree-types. Both factor would make any statistical inference questionable.

It is clear from the data that there is significant roll-off in graduate participation in their degreed field prior to the age of 30 and after 65 – neither is surprising. At the lower end, the age when someone enters the workforce fulltime after their undergrad graduation can vary considerably. Factors delaying may include entering graduate school, delayed hiring, or taking a break. Additionally, the roll-off around 65 can be attributed to retirement.

Therefore, the data will be subsetted to only consider ages 30-65. This should still give enough degrees of freedom perform an adequate linear fit and ensure the

```
dataTotalSubset <- filter(dataTotal, Age %in% 25:65)
dataGenderSubset <- filter(dataGender, Age %in% 25:65)

Mean <- dataTotalSubset %>% summarize(across(ArtsMatch:STEMMatch, ~ mean(.x, na.rm=TRUE)))
Median <- dataTotalSubset %>% summarize(across(ArtsMatch:STEMMatch, ~ median(.x, na.rm=TRUE)))
SD <- dataTotalSubset %>% summarize(across(ArtsMatch:STEMMatch, ~ sd(.x, na.rm=TRUE)))
statsTotal <- as.data.frame(cbind(t(Mean[1,]), t(Median[1,]), t(SD[1,])))
names(statsTotal) <- c("Mean", "Median", "SD")

MeanMale <- dataGenderSubset %>% filter(Gender==1) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                        ~ mean(.x, na.rm=TRUE)))
MedianMale <- dataGenderSubset %>% filter(Gender==1) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                          ~ median(.x, na.rm=TRUE)))
```

```
SDMale <- dataGenderSubset %>% filter(Gender==1) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                     ~ sd(.x, na.rm=TRUE)))
statsGender <- as.data.frame(cbind(t(MeanMale[1,]), t(MedianMale[1,]), t(SDMale[1,])))
names(statsGender) <- c("Mean", "Median", "SD")
row.names(statsGender) <- paste(row.names(statsGender), " (Male)")

MeanFemale <- dataGenderSubset %>% filter(Gender==2) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                    ~ mean(.x, na.rm=TRUE)))
MedianFemale <- dataGenderSubset %>% filter(Gender==2) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                      ~ median(.x, na.rm=TRUE)))
SDFemale <- dataGenderSubset %>% filter(Gender==2) %>% summarize(across(ArtsMatch:STEMMatch,
                                                                     ~ sd(.x, na.rm=TRUE)))
statsGender2 <- as.data.frame(cbind(t(MeanFemale[1,]), t(MedianFemale[1,]), t(SDFemale[1,])))
names(statsGender2) <- c("Mean", "Median", "SD")
row.names(statsGender2) <- paste(row.names(statsGender2), " (Female)")
statsGender <- rbind(statsGender, statsGender2)
```

Table 3: Summary statistics for the data without gender included for ages 30-65

Name	Mean	Median	SD
ArtsMatch	0.15543375	0.15032019	0.030182128
BusMatch	0.45333194	0.45923024	0.024462199
CompDirect	0.10812798	0.10941297	0.016255304
CompIndir	0.46463586	0.45594028	0.067160863
CompSTEM	0.51609894	0.51137865	0.066579681
EduMatch	0.57519622	0.60241407	0.096262139
EngrDirect	0.10531440	0.10268092	0.020885363
EngrIndir	0.27385818	0.26547753	0.030378464
EngrSTEM	0.46889538	0.47150509	0.068920694
HealthMatch	0.60485504	0.61375230	0.038898434
LifeSciDirect	0.01482577	0.01481851	0.004732834
LifeSciIndir	0.06195201	0.06161018	0.011871086
LifeSciSTEM	0.15315533	0.15393251	0.023548196
MathDirect	0.03397242	0.02713215	0.019688735
MathIndir	0.06459628	0.06593355	0.025900873
MathSTEM	0.27713840	0.27577737	0.050453005
PhySciDirect	0.06657772	0.06424493	0.013196367
PhySciIndir	0.09822434	0.10149415	0.018623285
PhySciSTEM	0.25544971	0.25714286	0.023601433
SocSciMatch	0.18152603	0.18110958	0.013457530
ANYMATCH	0.34347387	0.34871043	0.019964882
STEMtoSTEM	0.35871421	0.35685800	0.045673853

Name	Mean	Median	SD
STEMMatch	0.14173434	0.14237239	0.026337516

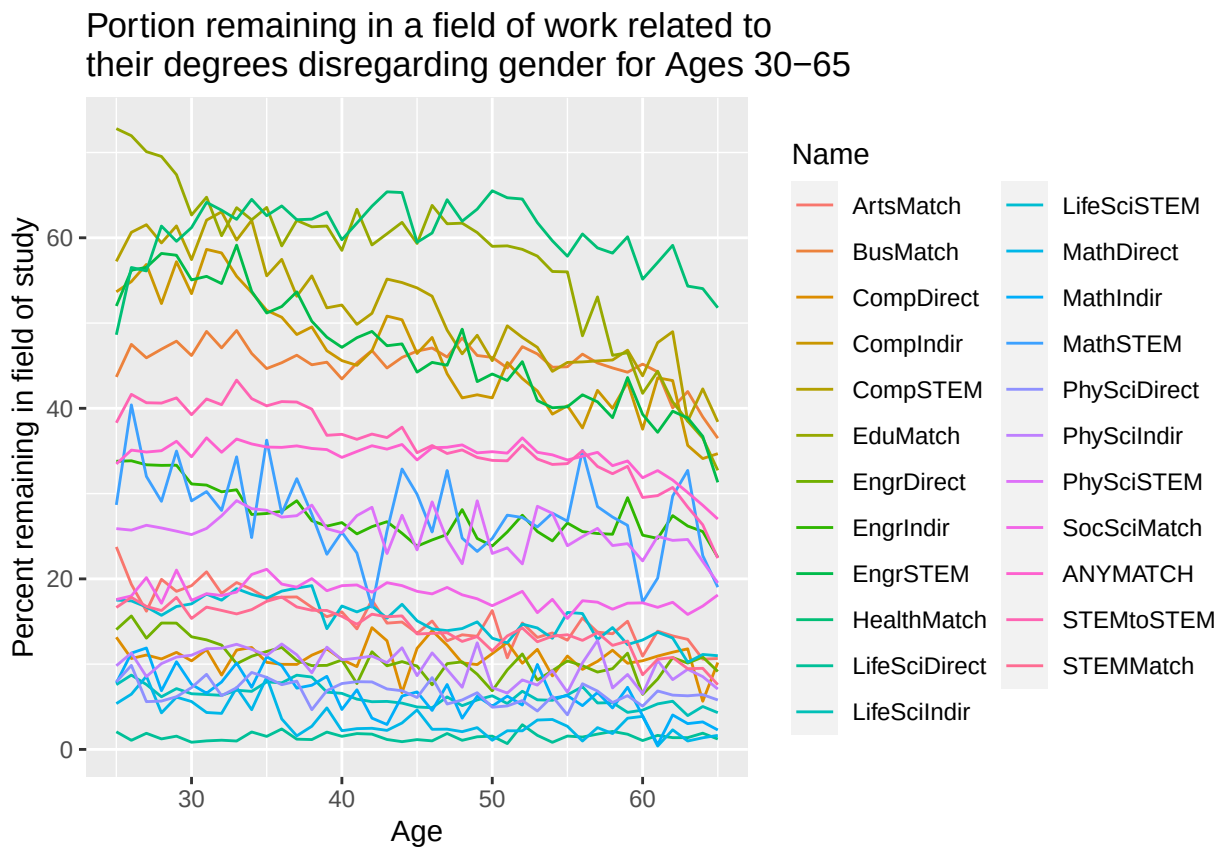
Table 4: Summary statistics for the data broken out by gender for ages 30-65

Name	Mean	Median	SD
ArtsMatch (Male)	0.18612428	0.18665179	0.034759389
BusMatch (Male)	0.48442148	0.48882863	0.025572500
CompDirect (Male)	0.11573149	0.11530320	0.019892466
CompIndir (Male)	0.51585247	0.50890017	0.067329563
CompSTEM (Male)	0.57486085	0.57099494	0.064725680
EduMatch (Male)	0.48847861	0.51068405	0.107649535
EngrDirect (Male)	0.11315252	0.10982420	0.023268962
EngrIndir (Male)	0.29145981	0.28217491	0.033660396
EngrSTEM (Male)	0.49291492	0.48819741	0.074504912
HealthMatch (Male)	0.55315233	0.56176355	0.062052321
LifeSciDirect (Male)	0.01700162	0.01712420	0.007110173
LifeSciIndir (Male)	0.06049585	0.05895517	0.013946881
LifeSciSTEM (Male)	0.17067299	0.17904550	0.034508228
MathDirect (Male)	0.03478948	0.02698242	0.025426886
MathIndir (Male)	0.06791252	0.06714899	0.032885533
MathSTEM (Male)	0.33669243	0.33764528	0.064702967
PhySciDirect (Male)	0.08020477	0.07742792	0.018844710
PhySciIndir (Male)	0.11394445	0.11379582	0.024355600
PhySciSTEM (Male)	0.31445917	0.31671716	0.034800410
SocSciMatch (Male)	0.13853131	0.14043350	0.019809161
ANYMATCH (Male)	0.32250110	0.32424482	0.019937538
STEMtoSTEM (Male)	0.42203994	0.41980669	0.061004567
STEMMatch (Male)	0.16979904	0.16734505	0.038381021
ArtsMatch (Female)	0.13803447	0.13150484	0.031236322
BusMatch (Female)	0.41940519	0.42490842	0.038920822
CompDirect (Female)	0.08806257	0.08552518	0.029325198
CompIndir (Female)	0.33737104	0.34166101	0.069529632
CompSTEM (Female)	0.37087412	0.37220697	0.073212320
EduMatch (Female)	0.59977669	0.62288216	0.094475648
EngrDirect (Female)	0.06974746	0.07015970	0.029079303
EngrIndir (Female)	0.19472607	0.18690024	0.044841569

Name	Mean	Median	SD
EngrSTEM (Female)	0.36188589	0.34473270	0.083306673
HealthMatch (Female)	0.61749868	0.63162704	0.041423481
LifeSciDirect (Female)	0.01287168	0.01151658	0.005566455
LifeSciIndir (Female)	0.06347949	0.06226760	0.018137115
LifeSciSTEM (Female)	0.13838548	0.13530492	0.025967008
MathDirect (Female)	0.03645855	0.03226562	0.025287609
MathIndir (Female)	0.06130336	0.05802669	0.038122967
MathSTEM (Female)	0.20866931	0.20265826	0.072766407
PhySciDirect (Female)	0.04779865	0.04518838	0.017917717
PhySciIndir (Female)	0.07583143	0.07575758	0.025154103
PhySciSTEM (Female)	0.17019476	0.17163678	0.043625665
SocSciMatch (Female)	0.21093653	0.21583638	0.022372887
ANYMATCH (Female)	0.36125383	0.36700072	0.021544734
STEMtoSTEM (Female)	0.23217059	0.23753560	0.039511777
STEMMatch (Female)	0.08759184	0.08806741	0.018293538

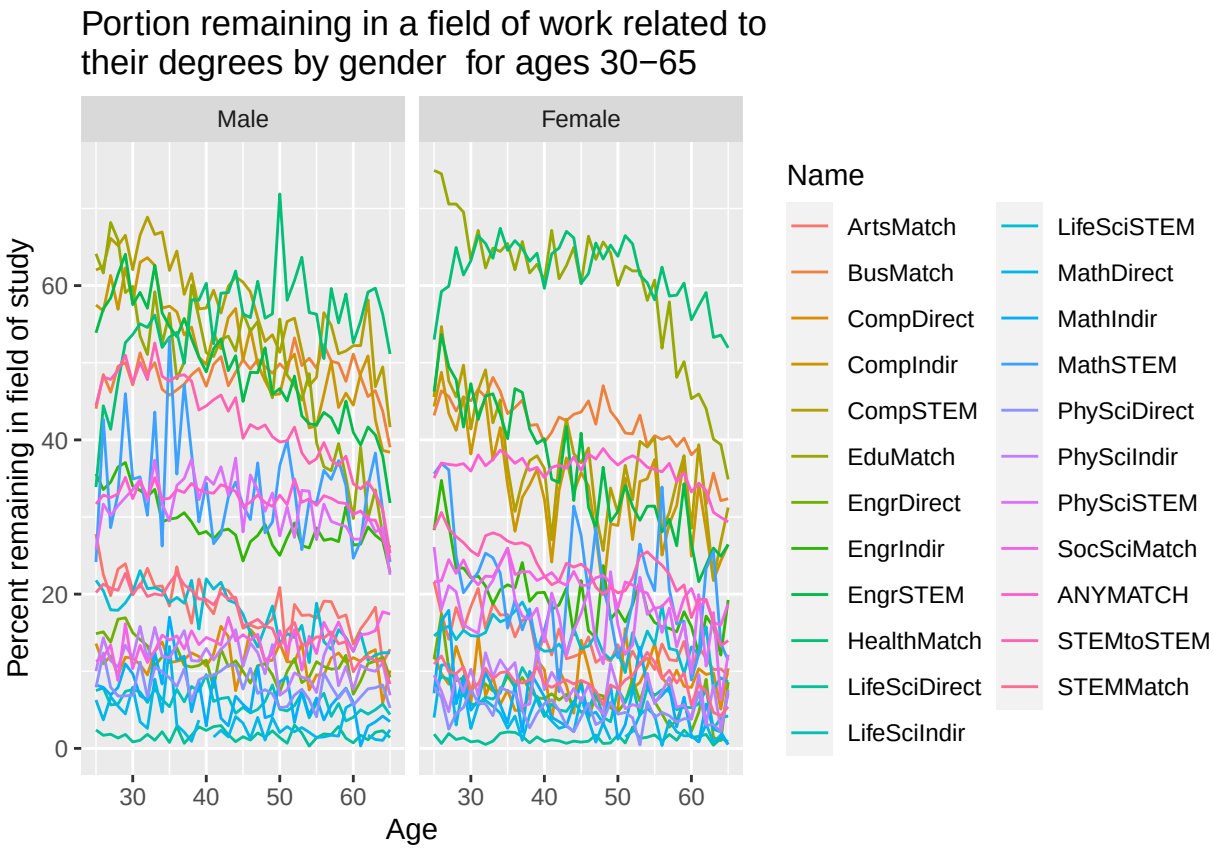
Graph All Degree-Types by Age for Ages 30-65

```
library(reshape2)
library(ggplot2)
meltTotalSubset <- melt(dataTotalSubset, id.vars="Age", variable.name="Name", value.name = "Percentage"
TotalSubsetPlot <- ggplot(meltTotalSubset, aes(x=Age, y=Percentage*100, group=Name))
TotalSubsetPlot <- TotalSubsetPlot +  geom_line(aes(color=Name))
TotalSubsetPlot <- TotalSubsetPlot +  labs(title="Portion remaining in a field of work related to\nthei
TotalSubsetPlot
```



Graph All Degree-Types by Age and by Gender for Ages 30-65

```
library(reshape2)
meltGenderSubset <- reshape2::melt(dataGenderSubset, id.vars=c("Age", "Gender"), variable.name="Name", v
GenderSubsetPlot <- ggplot(meltGenderSubset, aes(x=Age, y=Percentage*100, group=Name))
GenderSubsetPlot <- GenderSubsetPlot +  geom_line(aes(color=Name))
GenderSubsetPlot <- GenderSubsetPlot +  labs(title="Portion remaining in a field of work related to \nt
Gender.labs <- c("Male", "Female")
names(Gender.labs) <- c("1", "2")
GenderSubsetPlot <- GenderSubsetPlot +  facet_grid(. ~Gender, labeller = labeller(Gender = Gender.labs)
GenderSubsetPlot
```



Linear Regression Without Gender

The chief items of interest are to compare by degree-type:
1) the average involvement of people in a career matching their degree
2) the change in participation in a career matching their degree by age

Use a general linear model of the form $Y_{k,i} = \beta_{k,0} + \beta_{k,1} * X_{k,i} + \epsilon_{k,i}$, where:
* k : The k-th degree-type
* i : The i-th data point
* $Y_{k,i}$: The percent participation in the k-th degree-type at the i-th data point
* $X_{k,i}$: The age of the i-th data point for the k-th degree-type
* $\beta_{k,0}$: The intercept of the linear regressor for the k-th degree-type
* $\beta_{k,1}$: The slope of the linear regressor for the k-th degree-type showing change in percent participation per additional year in age.
* $\epsilon_{k,i}$: The uncharacterized noise in the system assumed to be $N\{0, \sigma^2\}$ (normally distributed with mean 0 and variance σ^2)

```
TotalSubsetFormula <- paste0("cbind(", paste(names(dataTotalSubset)[2:24], collapse = ", "), ")", " ~ Ag
fit <- lm(TotalSubsetFormula, dataTotalSubset)
fitcoeff <- coef(summary(fit))
# fitcoeff[[1]][1,4]  Here's how to pull out the p-value for the intercept for the first fit

TotalSubsetCoeff <- as.data.frame(fit$coefficients)
row.names(TotalSubsetCoeff) <- c("Beta0", "Beta1")
TotalSubsetCoeff <- t(TotalSubsetCoeff)
# Extract the t-values for all of the regressions
TotalCoeff <- coef(summary(fit))
```

```
TotalT <- data.frame()
for(i in 1:23) {
  tmp <- data.frame(Name = names(dataTotalSubset)[i+1],
                    Beta0T=TotalCoeff[[i]][1,3],
                    Beta1T=TotalCoeff[[i]][2,3])
  if(length(TotalT)==0) TotalT <- rbind(tmp)
  else TotalT <- rbind(TotalT, tmp)
}
# Extract the p-values for all of the regressions
TotalP <- data.frame()
for(i in 1:23) {
  tmp <- data.frame(Name = names(dataTotalSubset)[i+1],
                    Beta0P=TotalCoeff[[i]][1,4],
                    Beta0P95Sig=TotalCoeff[[i]][1,4]<0.025,
                    Beta1P=TotalCoeff[[i]][2,4],
                    Beta1P95Sig=TotalCoeff[[i]][2,4]<0.025,
                    AdjustedR2=summary(fit)[[i]]$adj.r.squared)
  if(length(TotalP)==0) TotalP <- rbind(tmp)
  else TotalP <- rbind(TotalP, tmp)
}

# Create the formula to pass to the linear regression for age and gender as predictors
GenderSubsetFormula <- paste0("cbind(", paste(names(dataGenderSubset)[3:25], collapse = ", "), ")", " ~
fit2 <- lm(GenderSubsetFormula, dataGenderSubset)

# Extract the coefficients for all of the regressions
GenderSubsetCoeff <- as.data.frame(fit2$coefficients)
row.names(GenderSubsetCoeff) <- c("Beta0", "Beta1","Beta2", "Beta3")
GenderSubsetCoeff <- t(GenderSubsetCoeff)
# Extract the t-values for all of the regressions
GenderCoeff <- coef(summary(fit2))
GenderT <- data.frame()
for(i in 1:23) {
  tmp <- data.frame(Name = names(dataGenderSubset)[i+2],
                    Beta0T=GenderCoeff[[i]][1,3],
                    Beta1T=GenderCoeff[[i]][2,3],
                    Beta2T=GenderCoeff[[i]][3,3],
                    Beta3T=GenderCoeff[[i]][4,3])
  if(length(GenderT)==0) GenderT <- rbind(tmp)
  else GenderT <- rbind(GenderT, tmp)
}
# Extract the p-values for all of the regressions
GenderP <- data.frame()
for(i in 1:23) {
  tmp <- data.frame(Name = names(dataGenderSubset)[i+2],
                    Beta0P=GenderCoeff[[i]][1,4],
                    Beta0Sig=GenderCoeff[[i]][1,4]<0.025,
                    Beta1P=GenderCoeff[[i]][2,4],
                    Beta1Sig=GenderCoeff[[i]][2,4]<0.025,
                    Beta2P=GenderCoeff[[i]][3,4],
                    Beta2Sig=GenderCoeff[[i]][3,4]<0.025,
                    Beta3P=GenderCoeff[[i]][4,4],
                    Beta3Sig=GenderCoeff[[i]][4,4]<0.025,
```

```
                    AdjustedR2=summary(fit2)[[i]]$adj.r.squared)
  if(length(GenderP)==0) GenderP <- rbind(tmp)
  else GenderP <- rbind(GenderP, tmp)
}
```

Table 5: Linear regression coefficients for degree-types disregarding gender for ages 30-65

Name	Beta0	Beta1
ArtsMatch	0.25220834	-0.00215054663
BusMatch	0.50638679	-0.00117899657
CompDirect	0.12053271	-0.00027566081
CompIndir	0.69526777	-0.00512515366
CompSTEM	0.74570952	-0.00510245722
EduMatch	0.89611384	-0.00713150266
EngrDirect	0.15754838	-0.00116075512
EngrIndir	0.35958116	-0.00190495507
EngrSTEM	0.71459768	-0.00546005093
HealthMatch	0.63844925	-0.00074653811
LifeSciDirect	0.01358195	0.00002764049
LifeSciIndir	0.09321559	-0.00069474632
LifeSciSTEM	0.22583247	-0.00161504755
MathDirect	0.08625048	-0.00116173457
MathIndir	0.13188962	-0.00149540772
MathSTEM	0.35807392	-0.00179856709
PhySciDirect	0.08890606	-0.00049618539
PhySciIndir	0.13528617	-0.00082359611
PhySciSTEM	0.29969026	-0.00098312334
SocSciMatch	0.21144871	-0.00066494834
ANYMATCH	0.39055139	-0.00104616716
STEMtoSTEM	0.51554332	-0.00348509139
STEMMatch	0.23121120	-0.00198837462

Table 6: Linear regression t-values for degree-types disregarding for gender for ages 30-65

Name	Beta0T	Beta1T
ArtsMatch	25.785840	-10.2305656
BusMatch	40.762762	-4.4159356
CompDirect	12.175603	-1.2956555

Name	Beta0T	Beta1T
CompIndir	41.058445	-14.0827099
CompSTEM	45.419071	-14.4602860
EduMatch	32.475412	-12.0254541
EngrDirect	16.254410	-5.5722018
EngrIndir	28.830587	-7.1067371
EngrSTEM	52.882476	-18.8007783
HealthMatch	27.115367	-1.4752650
LifeSciDirect	4.625254	0.4379738
LifeSciIndir	17.704429	-6.1397199
LifeSciSTEM	27.047054	-9.0001106
MathDirect	9.956763	-6.2401140
MathIndir	11.335335	-5.9801528
MathSTEM	12.619265	-2.9492865
PhySciDirect	12.132307	-3.1505396
PhySciIndir	13.770725	-3.9007370
PhySciSTEM	23.558208	-3.5958914
SocSciMatch	31.342084	-4.5860595
ANYMATCH	40.402881	-5.0357503
STEMtoSTEM	44.744536	-14.0740136
STEMMatch	33.075393	-13.2349807

Table 7: Linear regression p-values for degree-types disregarding for gender for ages 30-65

Name	Beta0P	Beta0P95Sig	Beta1P	Beta1P95Sig	AdjustedR2
ArtsMatch	4.1168e-26	TRUE	1.3365e-12	TRUE	0.72157347
BusMatch	1.3722e-33	TRUE	7.7477e-05	TRUE	0.31624502
CompDirect	7.3340e-15	TRUE	2.0271e-01	FALSE	0.01668497
CompIndir	1.0416e-33	TRUE	7.0338e-17	TRUE	0.83145314
CompSTEM	2.1996e-35	TRUE	2.9468e-17	TRUE	0.83877460
EduMatch	7.5637e-30	TRUE	1.0774e-14	TRUE	0.78214878
EngrDirect	5.8056e-19	TRUE	2.0334e-06	TRUE	0.42897468
EngrIndir	6.5501e-28	TRUE	1.5301e-08	TRUE	0.55310115
EngrSTEM	6.3944e-38	TRUE	3.7303e-21	TRUE	0.89808119
HealthMatch	6.4112e-27	TRUE	1.4817e-01	FALSE	0.02856992
LifeSciDirect	4.0542e-05	TRUE	6.6382e-01	FALSE	-0.02062111

Name	Beta0P	Beta0P95Sig	Beta1P	Beta1P95Sig	AdjustedR2
LifeSciIndir	3.0564e-20	TRUE	3.3126e-07	TRUE	0.47846151
LifeSciSTEM	7.0398e-27	TRUE	4.6301e-11	TRUE	0.66667219
MathDirect	2.8926e-12	TRUE	2.4032e-07	TRUE	0.48677827
MathIndir	6.5465e-14	TRUE	5.5183e-07	TRUE	0.46497046
MathSTEM	2.3928e-15	TRUE	5.3603e-03	TRUE	0.16139553
PhySciDirect	8.1918e-15	TRUE	3.1258e-03	TRUE	0.18243711
PhySciIndir	1.4610e-16	TRUE	3.6804e-04	TRUE	0.26220701
PhySciSTEM	1.1342e-24	TRUE	8.9719e-04	TRUE	0.22973879
SocSciMatch	2.8755e-29	TRUE	4.5793e-05	TRUE	0.33368806
ANYMATCH	1.9241e-33	TRUE	1.1187e-05	TRUE	0.37848419
STEMtoSTEM	3.9005e-35	TRUE	7.1775e-17	TRUE	0.83127906
STEMMatch	3.7955e-30	TRUE	5.2616e-16	TRUE	0.81322787

Linear Regression With Gender

The chief items of interest are to compare by degree-type:
1) the average involvement of men in a career matching their degree
2) the change in participation of men in a career matching their degree by age
3) the difference in the average involvement for women
4) the difference the slope of degree-type career by age for women

Use a general linear model of the form

$$Y_{k,i} = \beta_{k,0} + \beta_{k,1} * X_{1,k,i} + \beta_{k,2} * X_{2,k,i} + \beta_{k,3} * X_{1,k,i} * X_{2,k,i} + \epsilon_{k,i}$$

, where:
* k : The k-th degree-type
* i : The i-th data point
* $Y_{k,i}$: The percent participation in the k-th degree-type at the i-th data point
* $X_{1,k,i}$: The age of the i-th data point for the k-th degree-type
* $X_{2,k,i}$: The gender of the i-th data point for the k-th degree-type
* $\beta_{k,0}$: The intercept of the linear regressor for the k-th degree-type
* $\beta_{k,1}$: The slope of the linear regressor for the k-th degree-type showing change in percent participation per additional year in age.
* $\beta_{k,2}$: The change intercept of the linear regressor for the k-th degree-type due to gender
* $\beta_{k,3}$: The change in the slope of the linear regressor due to gender for the k-th degree-type showing change in percent participation per additional year in age.
* $\epsilon_{k,i}$: The uncharacterized noise in the system assumed to be $N\{0, \sigma^2\}$ (normally distributed with mean 0 and variance σ^2)

Table 8: Linear regression coefficients for degree-types accounting for gender for ages 30-65

Name	Beta0	Beta1	Beta2	Beta3
ArtsMatch	0.28373015	-0.00216552579	-0.050812142	0.00005311536
BusMatch	0.48963275	-0.00010663082	0.045811682	-0.00248224533
CompDirect	0.12961979	-0.00030360035	-0.041580045	0.00032476430

Name	Beta0	Beta1	Beta2	Beta3
CompIndir	0.74348588	-0.00506292978	-0.221733100	0.00098109687
CompSTEM	0.78838478	-0.00474252921	-0.224501143	0.00047802001
EduMatch	0.85530098	-0.00816076617	0.048556808	0.00138427852
EngrDirect	0.17064639	-0.00127171335	-0.029315628	-0.00031095938
EngrIndir	0.39050114	-0.00219362350	-0.070501116	-0.00058948159
EngrSTEM	0.75546454	-0.00581142037	-0.104626201	-0.00059716856
HealthMatch	0.43068231	0.00270955888	0.258732290	-0.00432357456
LifeSciDirect	0.01732379	-0.00001086145	-0.005678906	0.00003573078
LifeSciIndir	0.08869998	-0.00064492683	0.004848259	-0.00003247626
LifeSciSTEM	0.27447778	-0.00233710324	-0.080776599	0.00111211515
MathDirect	0.08108106	-0.00103519013	0.013682883	-0.00026409451
MathIndir	0.13257064	-0.00138644856	-0.002575527	-0.00011672168
MathSTEM	0.39207548	-0.00120382639	-0.075690239	-0.00114563440
PhySciDirect	0.10812493	-0.00062422869	-0.035653917	0.00007639280
PhySciIndir	0.16080081	-0.00105323858	-0.045302236	0.00016800614
PhySciSTEM	0.37035506	-0.00124666027	-0.116149252	-0.00062002980
SocSciMatch	0.10261913	0.00079496620	0.174051495	-0.00224372966
ANYMATCH	0.36916956	-0.00104078460	0.037267774	0.00003053410
STEMtoSTEM	0.63250600	-0.00467010403	-0.275642089	0.00190837570
STEMMatch	0.30226153	-0.00294358068	-0.182482790	0.00223199707

Table 9: Linear regression t-values for degree-types accounting for gender for ages 30-65

Name	Beta0T	Beta1T	Beta2T	Beta3T
ArtsMatch	20.894035	-7.3764168	-2.6711237	0.12940609
BusMatch	30.866849	-0.3109352	2.0616159	-5.17707298
CompDirect	8.163469	-0.8844452	-1.8693819	0.67668949
CompIndir	27.883790	-8.7830678	-5.9363470	1.21733108
CompSTEM	28.333457	-7.8838183	-5.7595635	0.56836216
EduMatch	27.541437	-12.1552480	1.1161625	1.47471756
EngrDirect	13.371287	-4.6092532	-1.6397776	-0.80611589
EngrIndir	23.208084	-6.0303751	-2.9910430	-1.15905747
EngrSTEM	40.241187	-14.3187215	-3.9783851	-1.05237705
HealthMatch	14.675161	4.2706078	6.2934117	-4.87401127
LifeSciDirect	4.237043	-0.1228778	-0.9915039	0.28912143
LifeSciIndir	10.186402	-3.4258863	0.3974594	-0.12339028

Name	Beta0T	Beta1T	Beta2T	Beta3T
LifeSciSTEM	20.854305	-8.2135647	-4.3811061	2.79547770
MathDirect	6.079818	-3.5905131	0.7324170	-0.65516132
MathIndir	6.954388	-3.3641906	-0.0964467	-0.20257267
MathSTEM	9.489343	-1.3477073	-1.3077237	-0.91733982
PhySciDirect	9.896019	-2.6426744	-2.3294404	0.23131557
PhySciIndir	11.375403	-3.4464389	-2.2877464	0.39320720
PhySciSTEM	16.811070	-2.6175227	-3.7635975	-0.93112371
SocSciMatch	10.212070	3.6593112	12.3643946	-7.38709517
ANYMATCH	34.127235	-4.4504235	2.4593387	0.09338521
STEMtoSTEM	43.805962	-14.9609917	-13.6277381	4.37270512
STEMMatch	30.836398	-13.8906308	-13.2896136	7.53341912

Landscape Table 10: Linear regression p-values for degree-types accounting for gender for ages 30-65

Name	Beta0P	Beta0Sig	Beta1P	Beta1Sig	Beta2P	Beta2Sig	Beta3P	Beta3Sig	AdjustedR2
ArtsMatch	5.3404e-33	TRUE	1.8174e-10	TRUE	9.2668e-03	TRUE	8.9738e-01	FALSE	0.72488644
BusMatch	2.2274e-44	TRUE	7.5671e-01	FALSE	4.2710e-02	FALSE	1.8285e-06	TRUE	0.70809222
CompDirect	5.8293e-12	TRUE	3.7928e-01	FALSE	6.5473e-02	FALSE	5.0068e-01	FALSE	0.20625174
CompIndir	2.4892e-41	TRUE	3.8507e-13	TRUE	8.4405e-08	TRUE	2.2729e-01	FALSE	0.85862365
CompSTEM	8.3095e-42	TRUE	1.9846e-11	TRUE	1.7502e-07	TRUE	5.7149e-01	FALSE	0.87111613
EduMatch	5.7958e-41	TRUE	2.0937e-19	TRUE	2.6792e-01	FALSE	1.4447e-01	FALSE	0.81670595
EngrDirect	1.5262e-21	TRUE	1.6257e-05	TRUE	1.0524e-01	FALSE	4.2272e-01	FALSE	0.64634317
EngrIndir	5.7436e-36	TRUE	5.7108e-08	TRUE	3.7603e-03	TRUE	2.5011e-01	FALSE	0.81928575
EngrSTEM	1.5837e-52	TRUE	3.7453e-23	TRUE	1.5865e-04	TRUE	2.9601e-01	FALSE	0.91576364
HealthMatch	9.5621e-24	TRUE	5.6437e-05	TRUE	1.8965e-08	TRUE	5.9528e-06	TRUE	0.43545483
LifeSciDirect	6.3677e-05	TRUE	9.0253e-01	FALSE	3.2463e-01	FALSE	7.7329e-01	FALSE	0.05795517
LifeSciIndir	8.5408e-16	TRUE	9.9728e-04	TRUE	6.9216e-01	FALSE	9.0213e-01	FALSE	0.23079839
LifeSciSTEM	6.0337e-33	TRUE	4.6798e-12	TRUE	3.7799e-05	TRUE	6.5777e-03	TRUE	0.62042931
MathDirect	4.6468e-08	TRUE	5.8649e-04	TRUE	4.6620e-01	FALSE	5.1437e-01	FALSE	0.28626257
MathIndir	1.1304e-09	TRUE	1.2121e-03	TRUE	9.2342e-01	FALSE	8.4002e-01	FALSE	0.23248110
MathSTEM	1.7555e-14	TRUE	1.8181e-01	FALSE	1.9496e-01	FALSE	3.6191e-01	FALSE	0.50495455
PhySciDirect	2.9975e-15	TRUE	1.0008e-02	TRUE	2.2530e-02	TRUE	8.1770e-01	FALSE	0.49804906
PhySciIndir	5.3803e-18	TRUE	9.3410e-04	TRUE	2.4972e-02	TRUE	6.9528e-01	FALSE	0.48486481
PhySciSTEM	3.7534e-27	TRUE	1.0708e-02	TRUE	3.3038e-04	TRUE	3.5478e-01	FALSE	0.81842498
SocSciMatch	7.6459e-16	TRUE	4.6775e-04	TRUE	8.8686e-20	TRUE	1.7349e-10	TRUE	0.85795946
ANYMATCH	1.9588e-47	TRUE	2.9318e-05	TRUE	1.6224e-02	TRUE	9.2585e-01	FALSE	0.63261935
STEMtoSTEM	3.5195e-55	TRUE	3.2372e-24	TRUE	5.5326e-22	TRUE	3.8976e-05	TRUE	0.95496982
STEMMatch	2.3857e-44	TRUE	1.9720e-22	TRUE	2.1121e-21	TRUE	9.1735e-11	TRUE	0.90626292

Linear Regression for gender based mean-center from ages 30-65

One problem with the previous analysis is that it is difficult to quantitatively determine the size of the effect gender has within the region of the data. The β_0 term is the intercept when $age = 0$ – not directly useful because it is outside of the 30 - 65 age range.

To shift the intercept to the useful age range, this analysis removes the mean of the combined male/female data for each degree-type. That means β_2 is the change in the average due to being female instead of male. For a well fit model, this should be nearly the same as the difference in means in Table 4.

```
# Remove the mean from each data column
center_colmeans <- function(x, xcenter) {
    x - rep(xcenter, rep.int(nrow(x), ncol(x)))
}
Means <-  colMeans(dataGenderSubset, na.rm = TRUE)
dataGenderSubsetCentered <- center_colmeans(dataGenderSubset, Means)

# Create the formula to pass to the linear regression for age and gender as predictors
GenderSubsetCenteredFormula <- paste0("cbind(", paste(names(dataGenderSubset)[3:25],
                                                      collapse = ", "), ")", " ~ Age * factor(Gender+.5)
fit3 <- lm(GenderSubsetCenteredFormula, dataGenderSubsetCentered)

# Extract the coefficients for all of the regressions
GenderSubsetCenteredCoeff <- as.data.frame(fit3$coefficients)
row.names(GenderSubsetCenteredCoeff) <- c("Beta0", "Beta1","Beta2", "Beta3")
GenderSubsetCenteredCoeff <- t(GenderSubsetCenteredCoeff)
# Extract the t-values for all of the regressions
GenderCenteredCoeff <- coef(summary(fit3))
GenderCenteredT <- data.frame()
for(i in 1:23) {
  tmp <- data.frame(Name = names(dataGenderSubsetCentered)[i+2],
                    Beta0T=GenderCenteredCoeff[[i]][1,3],
                    Beta1T=GenderCenteredCoeff[[i]][2,3],
                    Beta2T=GenderCenteredCoeff[[i]][3,3],
                    Beta3T=GenderCenteredCoeff[[i]][4,3])
  if(length(GenderCenteredT)==0) GenderCenteredT <- rbind(tmp)
  else GenderCenteredT <- rbind(GenderCenteredT, tmp)
}
# Extract the p-values for all of the regressions
GenderCenteredP <- data.frame()
for(i in 1:23) {
  tmp <- data.frame(Name = names(dataGenderSubsetCentered)[i+2],
                    Beta0P=GenderCenteredCoeff[[i]][1,4],
                    Beta0Sig=GenderCenteredCoeff[[i]][1,4]<0.025,
                    Beta1P=GenderCenteredCoeff[[i]][2,4],
                    Beta1Sig=GenderCenteredCoeff[[i]][2,4]<0.025,
                    Beta2P=GenderCenteredCoeff[[i]][3,4],
                    Beta2Sig=GenderCenteredCoeff[[i]][3,4]<0.025,
                    Beta3P=GenderCenteredCoeff[[i]][4,4],
                    Beta3Sig=GenderCenteredCoeff[[i]][4,4]<0.025,
                    AdjustedR2=summary(fit3)[[i]]$adj.r.squared)
  if(length(GenderCenteredP)==0) GenderCenteredP <- rbind(tmp)
  else GenderCenteredP <- rbind(GenderCenteredP, tmp)
}
```

Table 11: Linear regression coefficients for degree-types accounting for gender for ages 30-65 centered on the

means

Name	Beta0	Beta1	Beta2	Beta3
ArtsMatch	0.024202116	-0.00216552579	-0.048421950	0.00005311536
BusMatch	0.032921025	-0.00010663082	-0.065889358	-0.00248224533
CompDirect	0.014060747	-0.00030360035	-0.026965651	0.00032476430
CompIndir	0.089042285	-0.00506292978	-0.177583740	0.00098109687
CompSTEM	0.102103481	-0.00474252921	-0.202990242	0.00047802001
EduMatch	-0.056061148	-0.00816076617	0.110849341	0.00138427852
EngrDirect	0.021969304	-0.00127171335	-0.043308800	-0.00031095938
EngrIndir	0.048695141	-0.00219362350	-0.097027787	-0.00058948159
EngrSTEM	0.066550222	-0.00581142037	-0.131498786	-0.00059716856
HealthMatch	-0.032713045	0.00270955888	0.064171435	-0.00432357456
LifeSciDirect	0.001898370	-0.00001086145	-0.004071021	0.00003573078
LifeSciIndir	-0.002309396	-0.00064492683	0.003386828	-0.00003247626
LifeSciSTEM	0.014778896	-0.00233710324	-0.030731417	0.00111211515
MathDirect	-0.001137072	-0.00103519013	0.001798630	-0.00026409451
MathIndir	0.005572513	-0.00138644856	-0.007828003	-0.00011672168
MathSTEM	0.065222416	-0.00120382639	-0.127243787	-0.00114563440
PhySciDirect	0.016032927	-0.00062422869	-0.032216241	0.00007639280
PhySciIndir	0.018517137	-0.00105323858	-0.037741960	0.00016800614
PhySciSTEM	0.071928377	-0.00124666027	-0.144050593	-0.00062002980
SocSciMatch	-0.036341307	0.00079496620	0.073083660	-0.00224372966
ANYMATCH	-0.019543214	-0.00104078460	0.038641808	0.00003053410
STEMtoSTEM	0.095246053	-0.00467010403	-0.189765183	0.00190837570
STEMMatch	0.041104958	-0.00294358068	-0.082042922	0.00223199707

Table 12: Linear regression t-values for degree-types accounting for gender for ages 30-65 centered on the means

Name	Beta0T	Beta1T	Beta2T	Beta3T
ArtsMatch	6.9680487	-7.3764168	-9.9212361	0.12940609
BusMatch	8.1140225	-0.3109352	-11.5569649	-5.17707298
CompDirect	3.4622052	-0.8844452	-4.7252087	0.67668949
CompIndir	13.0561877	-8.7830678	-18.5305774	1.21733108
CompSTEM	14.3463987	-7.8838183	-20.2975292	0.56836216
EduMatch	-7.0578209	-12.1552480	9.9313302	1.47471756
EngrDirect	6.7302840	-4.6092532	-9.4418893	-0.80611589

Name	Beta0T	Beta1T	Beta2T	Beta3T
EngrIndir	11.3147151	-6.0303751	-16.0442636	-1.15905747
EngrSTEM	13.8594889	-14.3187215	-19.4888006	-1.05237705
HealthMatch	-4.3580059	4.2706078	6.0837907	-4.87401127
LifeSciDirect	1.8152731	-0.1228778	-2.7703210	0.28912143
LifeSciIndir	-1.0368994	-3.4258863	1.0821738	-0.12339028
LifeSciSTEM	4.3900711	-8.2135647	-6.4964754	2.79547770
MathDirect	-0.3333498	-3.5905131	0.3752491	-0.65516132
MathIndir	1.1428881	-3.3641906	-1.1425342	-0.20257267
MathSTEM	6.1716935	-1.3477073	-8.5685970	-0.91733982
PhySciDirect	5.7370484	-2.6426744	-8.2038219	0.23131557
PhySciIndir	5.1214556	-3.4464389	-7.4286464	0.39320720
PhySciSTEM	12.7649276	-2.6175227	-18.1927691	-0.93112371
SocSciMatch	-14.1392696	3.6593112	20.2354278	-7.38709517
ANYMATCH	-7.0633743	-4.4504235	9.9389233	0.09338521
STEMtoSTEM	25.7903106	-14.9609917	-36.5672061	4.37270512
STEMMatch	16.3951725	-13.8906308	-23.2878054	7.53341912

Table 13: Linear regression p-values for degree-types accounting for gender for ages 30-65 centered on the means

Name	Beta0P	Beta0Sig	Beta1P	Beta1Sig	Beta2P	Beta2Sig	Beta3P	Beta3Sig	AdjustedR2
ArtsMatch	1.0658e-09	TRUE	1.8174e-10	TRUE	2.6873e-15	TRUE	8.9738e-01	FALSE	0.72488644
BusMatch	7.2402e-12	TRUE	7.5671e-01	FALSE	2.5118e-18	TRUE	1.8285e-06	TRUE	0.70809222
CompDirect	8.8821e-04	TRUE	3.7928e-01	FALSE	1.0504e-05	TRUE	5.0068e-01	FALSE	0.20625174
CompIndir	5.3696e-21	TRUE	3.8507e-13	TRUE	1.0243e-29	TRUE	2.2729e-01	FALSE	0.85862365
CompSTEM	3.3666e-23	TRUE	1.9846e-11	TRUE	3.3968e-32	TRUE	5.7149e-01	FALSE	0.87111613
EduMatch	7.2339e-10	TRUE	2.0937e-19	TRUE	2.5723e-15	TRUE	1.4447e-01	FALSE	0.81670595
EngrDirect	2.9610e-09	TRUE	1.6257e-05	TRUE	2.1589e-14	TRUE	4.2272e-01	FALSE	0.64634317
EngrIndir	6.9457e-18	TRUE	5.7108e-08	TRUE	5.8677e-26	TRUE	2.5011e-01	FALSE	0.81928575
EngrSTEM	2.2273e-22	TRUE	3.7453e-23	TRUE	4.4365e-31	TRUE	2.9601e-01	FALSE	0.91576364
HealthMatch	4.1121e-05	TRUE	5.6437e-05	TRUE	4.5703e-08	TRUE	5.9528e-06	TRUE	0.43545483
LifeSciDirect	7.3480e-02	FALSE	9.0253e-01	FALSE	7.0557e-03	TRUE	7.7329e-01	FALSE	0.05795517
LifeSciIndir	3.0311e-01	FALSE	9.9728e-04	TRUE	2.8264e-01	FALSE	9.0213e-01	FALSE	0.23079839
LifeSciSTEM	3.6581e-05	TRUE	4.6798e-12	TRUE	8.0292e-09	TRUE	6.5777e-03	TRUE	0.62042931
MathDirect	7.3980e-01	FALSE	5.8649e-04	TRUE	7.0853e-01	FALSE	5.1437e-01	FALSE	0.28626257
MathIndir	2.5672e-01	FALSE	1.2121e-03	TRUE	2.5687e-01	FALSE	8.4002e-01	FALSE	0.23248110
MathSTEM	3.1637e-08	TRUE	1.8181e-01	FALSE	9.8617e-13	TRUE	3.6191e-01	FALSE	0.50495455
PhySciDirect	1.9195e-07	TRUE	1.0008e-02	TRUE	4.8840e-12	TRUE	8.1770e-01	FALSE	0.49804906
PhySciIndir	2.2759e-06	TRUE	9.3410e-04	TRUE	1.4480e-10	TRUE	6.9528e-01	FALSE	0.48486481
PhySciSTEM	1.7364e-20	TRUE	1.0708e-02	TRUE	3.1772e-29	TRUE	3.5478e-01	FALSE	0.81842498
SocSciMatch	7.4932e-23	TRUE	4.6775e-04	TRUE	4.1272e-32	TRUE	1.7349e-10	TRUE	0.85795946
ANYMATCH	7.0624e-10	TRUE	2.9318e-05	TRUE	2.4891e-15	TRUE	9.2585e-01	FALSE	0.63261935
STEMtoSTEM	5.0240e-39	TRUE	3.2372e-24	TRUE	1.4745e-49	TRUE	3.8976e-05	TRUE	0.95496982
STEMMatch	1.6523e-26	TRUE	1.9720e-22	TRUE	4.5806e-36	TRUE	9.1735e-11	TRUE	0.90626292

Sorting gender effects for evaluation

The tables below show the coefficients of linear regressions but sorted by the significance of the gender effects on either the average participation or the change in participation by age.

```
GenderSubsetCenteredCoeffSorted1 <- cbind(as.data.frame(GenderSubsetCenteredCoeff),
                                          Beta2P=GenderCenteredP[,6],
                                          Beta2Sig=GenderCenteredP[,7])
GenderSubsetCenteredCoeffSorted1 <-
  GenderSubsetCenteredCoeffSorted1[order(GenderSubsetCenteredCoeffSorted1$Beta2P),]

GenderSubsetCenteredCoeffSorted2 <- cbind(as.data.frame(GenderSubsetCenteredCoeff),
                                          Beta3P=GenderCenteredP[,8],
                                          Beta3Sig=GenderCenteredP[,9])
GenderSubsetCenteredCoeffSorted2 <-
  GenderSubsetCenteredCoeffSorted2[order(GenderSubsetCenteredCoeffSorted2$Beta3P),]
```

Table 14: Linear regression coefficients for degree-types accounting for gender for ages 30-65 centered on the means sorted by the significance of the effect of gender on the mean of the participation rate

Name	Beta0	Beta1	Beta2	Beta3	Beta2P	Beta2Sig
STEMtoSTEM	0.095246053	-0.00467010403	-0.189765183	0.00190837570	1.4745e-49	TRUE
STEMMatch	0.041104958	-0.00294358068	-0.082042922	0.00223199707	4.5806e-36	TRUE
CompSTEM	0.102103481	-0.00474252921	-0.202990242	0.00047802001	3.3968e-32	TRUE
SocSciMatch	-0.036341307	0.00079496620	0.073083660	-0.00224372966	4.1272e-32	TRUE
EngrSTEM	0.066550222	-0.00581142037	-0.131498786	-0.00059716856	4.4365e-31	TRUE
CompIndir	0.089042285	-0.00506292978	-0.177583740	0.00098109687	1.0243e-29	TRUE
PhySciSTEM	0.071928377	-0.00124666027	-0.144050593	-0.00062002980	3.1772e-29	TRUE
EngrIndir	0.048695141	-0.00219362350	-0.097027787	-0.00058948159	5.8677e-26	TRUE
BusMatch	0.032921025	-0.00010663082	-0.065889358	-0.00248224533	2.5118e-18	TRUE
ANYMATCH	-0.019543214	-0.00104078460	0.038641808	0.00003053410	2.4891e-15	TRUE
EduMatch	-0.056061148	-0.00816076617	0.110849341	0.00138427852	2.5723e-15	TRUE
ArtsMatch	0.024202116	-0.00216552579	-0.048421950	0.00005311536	2.6873e-15	TRUE
EngrDirect	0.021969304	-0.00127171335	-0.043308800	-0.00031095938	2.1589e-14	TRUE
MathSTEM	0.065222416	-0.00120382639	-0.127243787	-0.00114563440	9.8617e-13	TRUE
PhySciDirect	0.016032927	-0.00062422869	-0.032216241	0.00007639280	4.8840e-12	TRUE
PhySciIndir	0.018517137	-0.00105323858	-0.037741960	0.00016800614	1.4480e-10	TRUE
LifeSciSTEM	0.014778896	-0.00233710324	-0.030731417	0.00111211515	8.0292e-09	TRUE
HealthMatch	-0.032713045	0.00270955888	0.064171435	-0.00432357456	4.5703e-08	TRUE
CompDirect	0.014060747	-0.00030360035	-0.026965651	0.00032476430	1.0504e-05	TRUE
LifeSciDirect	0.001898370	-0.00001086145	-0.004071021	0.00003573078	7.0557e-03	TRUE
MathIndir	0.005572513	-0.00138644856	-0.007828003	-0.00011672168	2.5687e-01	FALSE
LifeSciIndir	-0.002309396	-0.00064492683	0.003386828	-0.00003247626	2.8264e-01	FALSE
MathDirect	-0.001137072	-0.00103519013	0.001798630	-0.00026409451	7.0853e-01	FALSE

Table 15: Linear regression coefficients for degree-types accounting for gender for ages 30-65 centered on the means sorted by the significance of the effect of gender on the slope of the participation rate by age

Name	Beta0	Beta1	Beta2	Beta3	Beta3P	Beta3Sig
STEMMatch	0.041104958	-0.00294358068	-0.082042922	0.00223199707	9.1735e-11	TRUE
SocSciMatch	-0.036341307	0.00079496620	0.073083660	-0.00224372966	1.7349e-10	TRUE
BusMatch	0.032921025	-0.00010663082	-0.065889358	-0.00248224533	1.8285e-06	TRUE
HealthMatch	-0.032713045	0.00270955888	0.064171435	-0.00432357456	5.9528e-06	TRUE
STEMtoSTEM	0.095246053	-0.00467010403	-0.189765183	0.00190837570	3.8976e-05	TRUE
LifeSciSTEM	0.014778896	-0.00233710324	-0.030731417	0.00111211515	6.5777e-03	TRUE
EduMatch	-0.056061148	-0.00816076617	0.110849341	0.00138427852	1.4447e-01	FALSE
CompIndir	0.089042285	-0.00506292978	-0.177583740	0.00098109687	2.2729e-01	FALSE
EngrIndir	0.048695141	-0.00219362350	-0.097027787	-0.00058948159	2.5011e-01	FALSE
EngrSTEM	0.066550222	-0.00581142037	-0.131498786	-0.00059716856	2.9601e-01	FALSE
PhySciSTEM	0.071928377	-0.00124666027	-0.144050593	-0.00062002980	3.5478e-01	FALSE
MathSTEM	0.065222416	-0.00120382639	-0.127243787	-0.00114563440	3.6191e-01	FALSE
EngrDirect	0.021969304	-0.00127171335	-0.043308800	-0.00031095938	4.2272e-01	FALSE
CompDirect	0.014060747	-0.00030360035	-0.026965651	0.00032476430	5.0068e-01	FALSE
MathDirect	-0.001137072	-0.00103519013	0.001798630	-0.00026409451	5.1437e-01	FALSE
CompSTEM	0.102103481	-0.00474252921	-0.202990242	0.00047802001	5.7149e-01	FALSE
PhySciIndir	0.018517137	-0.00105323858	-0.037741960	0.00016800614	6.9528e-01	FALSE
LifeSciDirect	0.001898370	-0.00001086145	-0.004071021	0.00003573078	7.7329e-01	FALSE
PhySciDirect	0.016032927	-0.00062422869	-0.032216241	0.00007639280	8.1770e-01	FALSE
MathIndir	0.005572513	-0.00138644856	-0.007828003	-0.00011672168	8.4002e-01	FALSE
ArtsMatch	0.024202116	-0.00216552579	-0.048421950	0.00005311536	8.9738e-01	FALSE
LifeSciIndir	-0.002309396	-0.00064492683	0.003386828	-0.00003247626	9.0213e-01	FALSE
ANYMATCH	-0.019543214	-0.00104078460	0.038641808	0.00003053410	9.2585e-01	FALSE