UX and MARS paper analysis output

Peter Sperounes, Ponpanich, Jul Applolnik, Ivo Wenzel March Sagan, Dominic Ernstein of Ap Apolnik and Sharon Ponpanich, Jul Applolnik, Ivo Wenzel March Sagan, Dominic Ernstein of Ap Apolnik and Sharon

```
ggsave("C:/Users/ponpanich/Desktop/Sperounes_Ponpanich_Applolnik_Wenzel_Sagan_Ernstein_Apolnik_Sharon_UX_and_MARS_paper_analysis_output_variable_distributions.pdf",
plot = ggarrange(ggplot(data=ux_mars_df, aes(x=Age)) + geom_density(),
ggplot(data=ux_mars_df, aes(x=Gender)) + geom_bar(),
ggplot(data=ux_mars_df, aes(x=Ethnicity)) + geom_bar(),
ggplot(data=ux_mars_df, aes(x=Education)) + geom_bar(),
ncol=2, nrow=2))
```

variable distributions

```
fig.height=7 ggplot(data=ux_mars_df,aes(x=Age)) +
  geom_density(fill="#ff0000") + ggtitle("Age") +
  ggplot(data=ux_mars_df,aes(x=Gender)) + geom_bar() +
  ggtitle("Gender") +
  ggplot(data=ux_mars_df,aes(x=Ethnicity)) + geom_bar() + ggtitle("Ethnicity") +
  ggplot(data=ux_mars_df,aes(x=Education)) + geom_bar() + ggtitle("Education") +
  ncol=2, nrow=2)
```

```
## Warning: Removed 2 rows containing non-finite values (stat_density).
```

Distribution of variables (n=50-56)

Figure 1.

```
## Warning: Removed 2 rows containing non-finite values (stat_density).
```

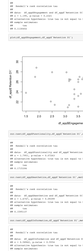
```
## ,Age,Gender,
## Female: 29.0000000 Female:0.5000000,
## Male: 29.0000000 Male:0.5000000,
## Caucasian: 29.0000000 Caucasian:0.5000000,
## Hispanic: 29.0000000 Hispanic:0.5000000,
## African American: 29.0000000 African American:0.5000000,
## Asian: 29.0000000 Asian:0.5000000,
## Other: 29.0000000 Other:0.5000000,
## Bachelor's degree: 29.0000000 Bachelor's degree:0.5000000,
## Master's degree: 29.0000000 Master's degree:0.5000000,
## Doctorate: 29.0000000 Doctorate:0.5000000,
## Some college: 29.0000000 Some college:0.5000000,
## High school: 29.0000000 High school:0.5000000,
## Vocational degree: 29.0000000 Vocational degree:0.5000000,
## Other: 29.0000000 Other:0.5000000,
## Density: 0.0000000 Density:0.0000000,
## count: 1.0000000 count:1.0000000,
## Ethnicity,Education,
## Caucasian: 29.0000000 Bachelor's degree: 29.0000000,
## Hispanic: 29.0000000 Master's degree: 29.0000000,
## African American: 29.0000000 Doctorate: 29.0000000,
## Asian: 29.0000000 Some college: 29.0000000,
## Other: 29.0000000 High school: 29.0000000,
## Density: 0.0000000 Vocational degree: 29.0000000,
## count: 1.0000000 Other: 29.0000000,
## Density: 0.0000000,
## count: 1.0000000,
```

Pre-registered analyses

```
fit.lm.age=lm(Age ~ Gender + Ethnicity + Education + Gender:Ethnicity + Gender:Education,
data=ux_mars_df)
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age,y=Age)) + geom_jitter(height=0.1)
```

```
##
## Call:
## lm(formula = Age ~ Gender + Ethnicity + Education + Gender:Ethnicity +
## Gender:Education, data = ux_mars_df)
##
## Coefficients:
## (Intercept)       Gender  Ethnicity  Education  Gender:Ethnicity
##  -1.0000000  -0.0000000   0.0000000   0.0000000   0.0000000
## Gender:Education
##   0.0000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age,y=Age)) + geom_jitter(height=0.1)
```




```
fit.lm.age.gender=lm(Age ~ Gender,data=ux_mars_df)
```

```
##
## Call:
## lm(formula = Age ~ Gender, data = ux_mars_df)
##
## Coefficients:
## (Intercept)       Gender
##  -1.0000000   0.0000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age.gender,y=Age)) +
  geom_jitter(height=0.1) + ylab("Age") +
  ggtitle("NAMS Mean Predicting Age Revenue") +facet_wrap(~ Gender)
```

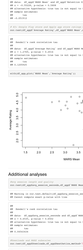
```
## Warning: Removed 2 rows containing non-finite values (stat_smooth).
```

```
fit.lm.age.ethnicity=lm(Age ~ Ethnicity,data=ux_mars_df)
```

```
##
## Call:
## lm(formula = Age ~ Ethnicity, data = ux_mars_df)
##
## Coefficients:
## (Intercept)  Ethnicity
##  -1.0000000   0.0000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age.ethnicity,y=Age)) +
  geom_jitter(height=0.1) + ylab("Age") +
  ggtitle("NAMS Mean Predicting Age Revenue") +facet_wrap(~ Ethnicity)
```

```
## Warning: Removed 2 rows containing non-finite values (stat_smooth).
```




```
fit.lm.age.education=lm(Age ~ Education,data=ux_mars_df)
```

```
##
## Call:
## lm(formula = Age ~ Education, data = ux_mars_df)
##
## Coefficients:
## (Intercept)  Education
##  -1.0000000   0.0000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age.education,y=Age)) +
  geom_jitter(height=0.1) + ylab("Age") +
  ggtitle("NAMS Mean Predicting Age Revenue") +facet_wrap(~ Education)
```

```
## Warning: Removed 2 rows containing non-finite values (stat_smooth).
```

```
fit.lm.age.gender.ethnicity=lm(Age ~ Gender + Ethnicity,data=ux_mars_df)
```

```
##
## Call:
## lm(formula = Age ~ Gender + Ethnicity, data = ux_mars_df)
##
## Coefficients:
## (Intercept)       Gender  Ethnicity
##  -1.0000000   0.0000000   0.0000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age.gender.ethnicity,y=Age)) +
  geom_jitter(height=0.1) + ylab("Age") +
  ggtitle("NAMS Mean Predicting Age Revenue") +facet_wrap(~ Gender,Ethnicity)
```

```
## Warning: Removed 2 rows containing non-finite values (stat_smooth).
```

```
fit.lm.age.gender.education=lm(Age ~ Gender + Education,data=ux_mars_df)
```

```
##
## Call:
## lm(formula = Age ~ Gender + Education, data = ux_mars_df)
##
## Coefficients:
## (Intercept)       Gender  Education
##  -1.0000000   0.0000000   0.0000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age.gender.education,y=Age)) +
  geom_jitter(height=0.1) + ylab("Age") +
  ggtitle("NAMS Mean Predicting Age Revenue") +facet_wrap(~ Gender,Education)
```

```
## Warning: Removed 2 rows containing non-finite values (stat_smooth).
```



```
fit.lm.age.gender.ethnicity.education=lm(Age ~ Gender + Ethnicity + Education,
data=ux_mars_df)
```

```
##
## Call:
## lm(formula = Age ~ Gender + Ethnicity + Education, data = ux_mars_df)
##
## Coefficients:
## (Intercept)       Gender  Ethnicity  Education
##  -1.0000000   0.0000000   0.0000000   0.0000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
ggplot(data=ux_mars_df,aes(x=fitted.fit.lm.age.gender.ethnicity.education,y=Age)) +
  geom_jitter(height=0.1) + ylab("Age") +
  ggtitle("NAMS Mean Predicting Age Revenue") +facet_wrap(~ Gender,Ethnicity,Education)
```

```
## Warning: Removed 2 rows containing non-finite values (stat_smooth).
```

Pearson Chi-Squared and GAM p-values corrected 0.05, all ns randomness ns
Pearson Chi-Squared and GAM p-values corrected 0.05, all ns randomness ns

```
ANOVA GAM p.adjusted(p.fit,method="fdr",n=length(p.fit)) p.adjusted(p.fit,method="fdr",n=length(p.fit))
fit.lm.age.gender.ethnicity.education.pvalues=c(p.fit.lm.age,p.fit.lm.age.gender,p.fit.lm.age.ethnicity,p.fit.lm.age.education,p.fit.lm.age.gender.ethnicity,p.fit.lm.age.gender.education,p.fit.lm.age.ethnicity.education,p.fit.lm.age.gender.ethnicity.education)
```

```
##
## Call:
## lm(formula = fit.lm.age.gender.ethnicity.education.pvalues ~
## 1)
##
## Coefficients:
## (Intercept)
##  -1.000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

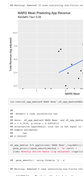
```
plot(fit.lm.age.gender.ethnicity.education.pvalues)
```

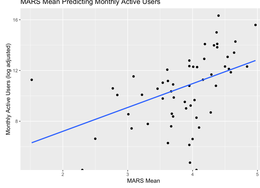


```
fit.lm.age.gender.ethnicity.education.pvalues.fdr=p.adjust(fit.lm.age.gender.ethnicity.education.pvalues,method="fdr",n=length(fit.lm.age.gender.ethnicity.education.pvalues))
```

```
##
## Call:
## lm(formula = fit.lm.age.gender.ethnicity.education.pvalues.fdr ~
## 1)
##
## Coefficients:
## (Intercept)
##  -1.000000
##
## NA degrees of freedom
## NA residual
## NA residual sum of squares
## NA R-squared, NA adjusted R-squared
## NA F-statistic, NA DF, NA DF, NA p-value
```

```
plot(fit.lm.age.gender.ethnicity.education.pvalues.fdr)
```


fit.lm.age.gender.ethnicity.education.pvalues.fdr



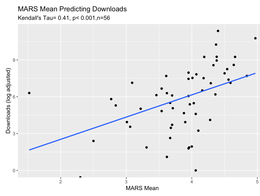

```
cor.test(df_app_median$`MARS Mean`,df_app_median$Downloads,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$`MARS Mean` and df_app_median$Downloads
## z = 4.4331, p-value = 0.00009929
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.4068712
```

```
df_app_median %>% ggplot(aes(`MARS Mean`,log(Downloads)))+
  geom_point()+geom_smooth(method = 'lm',se=F)+
  ylab('Downloads (log adjusted)')+
  ggtitle('MARS Mean Predicting Downloads',"Kendall's Tau= 0.41, p< 0.001,n=56")
```

```
## `geom_smooth()` using formula `y ~ x`
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

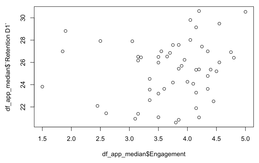



```
#note that Downloads and MAD are correlated 0.99, so its redundant to
#include both in the same paper. Revenue is correlated with them ~.45


#RQ:The MARS's Engagement, Functionality, and Aesthetics subscales will predict
#user retention more strongly than the Information subscale
cor.test(df_app_median$Engagement,df_app_median$`Retention D1`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Engagement and df_app_median$`Retention D1`
## z = 1.195, p-value = 0.2321
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.1130402
```

```
plot(df_app_median$Engagement,df_app_median$`Retention D1`)
```




```
cor.test(df_app_median$Functionality,df_app_median$`Retention D1`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Functionality and df_app_median$`Retention D1`
## z = 1.7951, p-value = 0.07263
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.1715306
```

```
cor.test(df_app_median$Aesthetics,df_app_median$`Retention D1`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Aesthetics and df_app_median$`Retention D1`
## z = 1.6747, p-value = 0.09399
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.1595137
```

```
cor.test(df_app_median$Information,df_app_median$`Retention D1`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Information and df_app_median$`Retention D1`
## z = 0.58962, p-value = 0.5556
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.05567802
```

```
cor.test(df_app_median$`MARS Mean`,df_app_median$`Retention D1`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$`MARS Mean` and df_app_median$`Retention D1`
## z = 1.6942, p-value = 0.09023
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.1592992
```

```
cor.test(df_app_median$Engagement,df_app_median$`Retention D7`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Engagement and df_app_median$`Retention D7`
## z = -0.26887, p-value = 0.788
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.02543604
```

```
cor.test(df_app_median$Functionality,df_app_median$`Retention D7`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Functionality and df_app_median$`Retention D7`
## z = 0.27675, p-value = 0.782
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.02644429
```

```
cor.test(df_app_median$Aesthetics,df_app_median$`Retention D7`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Aesthetics and df_app_median$`Retention D7`
## z = 0.089717, p-value = 0.9285
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.008545574
```

```
cor.test(df_app_median$Information,df_app_median$`Retention D7`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Information and df_app_median$`Retention D7`
## z = -0.6344, p-value = 0.5258
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.05969154
```

```
cor.test(df_app_median$`MARS Mean`,df_app_median$`Retention D7`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$`MARS Mean` and df_app_median$`Retention D7`
## z = -0.097025, p-value = 0.9227
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.009122863
```

```
cor.test(df_app_median$Engagement,df_app_median$`Retention D30`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Engagement and df_app_median$`Retention D30`
## z = -0.60494, p-value = 0.5453
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.0572066
```

```
cor.test(df_app_median$Functionality,df_app_median$`Retention D30`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Functionality and df_app_median$`Retention D30`
## z = -0.32142, p-value = 0.7477
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.03072182
```

```
cor.test(df_app_median$Aesthetics,df_app_median$`Retention D30`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Aesthetics and df_app_median$`Retention D30`
## z = -0.35138, p-value = 0.7253
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.03345769
```

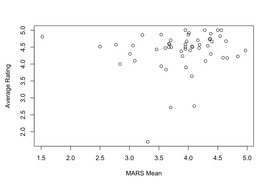
```
cor.test(df_app_median$Information,df_app_median$`Retention D30`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$Information and df_app_median$`Retention D30`
## z = -1.1041, p-value = 0.2693
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.1038972
```

```
cor.test(df_app_median$`MARS Mean`,df_app_median$`Retention D30`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$`MARS Mean` and df_app_median$`Retention D30`
## z = -0.55228, p-value = 0.5808
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## -0.051912
```

```
# H1: Google Play store and Apple app store ratings will be correlated with the MARS mean.
cor.test(df_app_median$`Average Rating`,df_app_median$`MARS Mean`,method='kendall')
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_app_median$`Average Rating` and df_app_median$`MARS Mean`
## z = 1.2769, p-value = 0.2016
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.1205925
```

```
with(df_app_median,plot(`MARS Mean`,`Average Rating`))
```