

```

/*
=====
=====
Do file name:   loyalty_missing_do
Project:       The Impact of Missing Data on the Predictability of
Customer Loyalty Models
Date:         29/03/2021
Author:       Dominik Piehlmaier
Description:   Web appendix to book chapter
=====
=====*/

```

```

*Prep
version 16.1
clear
graph close _all
set more off
cap ssc install outreg2
cap ssc install univar
pwd
*cd

*Import and Prepare Data
import delimited "https://www.neuraldesigner.com/files/datasets/
bank_churn.csv", clear
destring *, replace
ren gender gender_string
ren country country_string
encode gender_string, gen(gender)
encode country_string, gen(country)
gen male=(gender==2)
gen germany=(country==2)
gen spain=(country==3)

*Full Model Without Missing Data
reg churn credit_score i.country male age tenure balance products
credit_card active_member estimated_salary, robust
outreg2 using MCAR.doc, label ctitle(Complete Model) alpha(0.001,
0.01, 0.05) dec(2)

*MCAR based on ID
gen id=customer_id/1000000
univar id
replace country=. if id>15.75
replace male=. if id>15.75
*-Listwise
reg churn credit_score i.country male age tenure balance products
credit_card active_member estimated_salary, robust
outreg2 using MCAR.doc, append label ctitle(Listwise Model)
alpha(0.001, 0.01, 0.05) dec(2)
*-FIML
sem (churn <- credit_score germany spain male age tenure balance
products credit_card active_member estimated_salary), method(mlmv)

```

```

vce(robust) nocapslatent
estat gof, stats(all)
*-PMM
mi set mlong
mi register imputed country gender
mi impute chained (pmm, knn(5)) country gender = churn credit_score
age tenure balance products credit_card active_member
estimated_salary, add(10)

mi estimate, post: reg churn credit_score i.country i.gender age
tenure balance products credit_card active_member estimated_salary,
robust
outreg2 using MCAR.doc, append label ctitle(PMM Model) alpha(0.001,
0.01, 0.05) dec(2)
*repeat with mibeta (findit mibeta) for  $\hat{R}^2$ 

```

```

*Second Run: MAR – Import and Prepare Data
import delimited "https://www.neuraldesigner.com/files/datasets/
bank_churn.csv", clear
destring *, replace
ren gender gender_string
ren country country_string
encode gender_string, gen(gender)
encode country_string, gen(country)
gen male=(gender==2)
gen germany=(country==2)
gen spain=(country==3)

```

```

*Full Model Without Missing Data
reg churn credit_score germany spain male age tenure balance
products credit_card active_member estimated_salary, robust
outreg2 using MAR.doc, label ctitle(Complete Model) alpha(0.001,
0.01, 0.05) dec(2)

```

```

*MAR based on Country
replace age=. if germany==1
replace male=. if germany==1
*-Listwise
reg churn credit_score germany spain male age tenure balance
products credit_card active_member estimated_salary, robust
outreg2 using MAR.doc, append label ctitle(Listwise Model)
alpha(0.001, 0.01, 0.05) dec(2)
*-FIML
sem (churn <- credit_score germany spain male age tenure balance
products credit_card active_member estimated_salary), method(mlmv)
vce(robust) nocapslatent
estat gof, stats(all)
outreg2 using MAR.doc, append label ctitle(FIML Model) alpha(0.001,
0.01, 0.05) dec(2)
*-PMM
mi set mlong
mi register imputed age male
mi impute chained (pmm, knn(5)) age male = churn credit_score

```

```
germany spain tenure balance products credit_card active_member
estimated_salary, add(10)
```

```
mi estimate, post: reg churn credit_score germany spain male age
tenure balance products credit_card active_member estimated_salary,
robust
```

```
outreg2 using MNAR.doc, append label ctitle(PMM Model) alpha(0.001,
0.01, 0.05) dec(2)
```

```
*repeat with mibeta (findit mibeta) for  $\tilde{R}, \hat{A}^2$ 
```

```
*Third Run: MNAR – Import and Prepare Data
```

```
import delimited "https://www.neuraldesigner.com/files/datasets/
bank_churn.csv", clear
```

```
destring *, replace
```

```
ren gender gender_string
```

```
ren country country_string
```

```
encode gender_string, gen(gender)
```

```
encode country_string, gen(country)
```

```
gen male=(gender==2)
```

```
gen germany=(country==2)
```

```
gen spain=(country==3)
```

```
*Full Model Without Missing Data
```

```
reg churn credit_score germany spain male age tenure balance
```

```
products credit_card active_member estimated_salary, robust
```

```
outreg2 using MNAR.doc, label ctitle(Complete Model) alpha(0.001,
0.01, 0.05) dec(2)
```

```
*MNAR based on Country
```

```
replace churn=. if churn==0 & germany==1
```

```
replace active_member=. if active_member==0 & male==0
```

```
*-Listwise
```

```
reg churn credit_score germany spain male age tenure balance
```

```
products credit_card active_member estimated_salary, robust
```

```
outreg2 using MNAR.doc, append label ctitle(Listwise Model)
```

```
alpha(0.001, 0.01, 0.05) dec(2)
```

```
*-FIML
```

```
sem (churn <- credit_score germany spain male age tenure balance
```

```
products credit_card active_member estimated_salary), method(mlmv)
```

```
vce(robust) nocapslatent
```

```
estat gof, stats(all)
```

```
outreg2 using MNAR.doc, append label ctitle(FIML Model) alpha(0.001,
0.01, 0.05) dec(2)
```

```
*-PMM
```

```
mi set mlong
```

```
mi register imputed active_member churn
```

```
mi impute chained (pmm, knn(5)) active_member churn = male
```

```
credit_score germany spain tenure balance products credit_card age
```

```
estimated_salary, add(10)
```

```
mi estimate, post: reg churn credit_score germany spain male age
tenure balance products credit_card active_member estimated_salary,
robust
```

```
outreg2 using MNAR.doc, append label ctitle(PMM Model) alpha(0.001,
```

0.01, 0.05) dec(2)

*repeat with mibeta (findit mibeta) for RÃ