1 Log files from Stata (external links)

SHARE-1-harmon (reshape data)
SHARE-2-harmon-Part5 (generation of disease variables)

2 Sample Selection

```
. tab hacohort wave,m

     hacohort: Sample |                                Survey Wave
               cohort | 2004 wave  2006/07 w  2011/12 w  2013 wave  2015 wave  2017 wave  2019/20 w |     Total
----------------------+-------------------------------------------------------------------------------+----------
1.Original sample for |    51,312     51,312     51,312     51,312     51,312     51,312     51,312 |   359,184
2.2006 Refreshment sa |     5,836      5,836      5,836      5,836      5,836      5,836      5,836 |    40,852
----------------------+-------------------------------------------------------------------------------+----------
                Total |    57,148     57,148     57,148     57,148     57,148     57,148     57,148 |   400,036

. tab iwstatr wave,m

                      |                                Survey Wave
   r interview status | 2004 wave  2006/07 w  2011/12 w  2013 wave  2015 wave  2017 wave  2019/20 w |     Total
----------------------+-------------------------------------------------------------------------------+----------
              0.inap. |    37,079     25,054     17,649     20,009     14,567      1,898      1,730 |   117,986
        1.resp, alive |    20,069     25,176     26,853     22,511     23,322     32,289     19,780 |   170,000
          4.nr, alive |         0      4,828      7,138      6,779      8,805      8,662     10,489 |    46,701
   5.nr, died this wv |         0        279        468        593        603        958      1,061 |     3,962
   6.nr, died prev wv |         0          0        660      1,096      1,895      2,799      3,983 |    10,433
9.nr, dk if alive or  |         0      1,811      4,380      6,160      7,956     10,542     20,105 |    50,954
----------------------+-------------------------------------------------------------------------------+----------
                Total |    57,148     57,148     57,148     57,148     57,148     57,148     57,148 |   400,036

. sum agemin,

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
      agemin |    400,036    59.13131    6.417423         14         70

. qui log close log
```

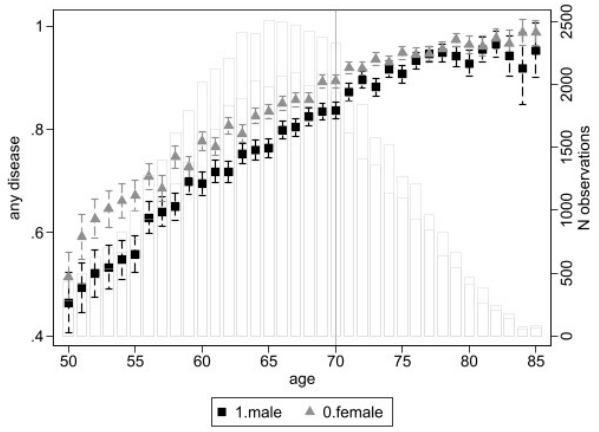
3 Choice of Diseases

```
. codebook d_* timetonextdisease2, compact

Variable         Obs Unique      Mean  Min  Max  Label
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------
d_hibp        169620      2  .4831859    0    1  ever had | taking meds for hibp
d_diab        169598      2  .1500018    0    1  ever had | taking meds for diab
d_heart       169596      2  .2217269    0    1  ever had | taking meds for heart
d_lung        169587      2  .0859913    0    1  ever had | taking meds for lung
d_psych       169856      2  .2651246    0    1  ever had | taking meds for psych
d_osteo       169583      2  .0953456    0    1  ever had | taking meds for osteo
d_cancr       169328      2   .082727    0    1  (only) ever had cancr
d_strok       169334      2  .0539053    0    1  (only) ever had strok
d_arthr       169406      2  .3204373    0    1  (only) ever had arthr
d_any         169929      2  .7559922    0    1  any disease
d_miss        169929      9  .0203203    0    8  # miss.diseases
d_count       169261     10  1.758054    0    9  # diseases
d_count_geq2  169261      2  .4889727    0    1  >=2 diseases
d_count_in~x  169261     10  .1953394    0    1  disease index (=count/total diseases)
d_anyatfir~s  399700      2  .6346585    0    1  already has disease at baseline
d_anyever     399700      2  .7874256    0    1  ever reports any disease
d_anyever_g2  400036      2  .4975852    0    1  ever reports having had any disease (g2aging)
timetonext~2   53524    179  54.78008   11  190  time (months) from C to C+1 (or more) diseases
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------

. codebook diff_*,compact // assuming the first disease starts with hibp in the dataset

Variable         Obs Unique      Mean  Min  Max  Label
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------
diff_d_count   85319      8  .3377091   -1    6  1st diff of # of diseases
diff_miss~nt  110881     11  .3521162   -3    7  1st diff of # of diseases: (L(t-2) used if L(t-1) missing)
diff_d_hibp    85632      2  .0674631    0    1  1st diff of d_hibp ('ever had' | medication)
diff_d_diab    85628      2  .0281684    0    1  1st diff of d_diab ('ever had' | medication)
diff_d_heart   85627      2  .0567344    0    1  1st diff of d_heart ('ever had' | medication)
diff_d_lung    85626      2  .0220494    0    1  1st diff of d_lung ('ever had' | medication)
diff_d_psych   85656      3  .0203488   -1    1  1st diff of d_psych ('ever had' | medication)
diff_d_osteo   85626      2  .0325485    0    1  1st diff of d_osteo ('ever had' | medication)
```

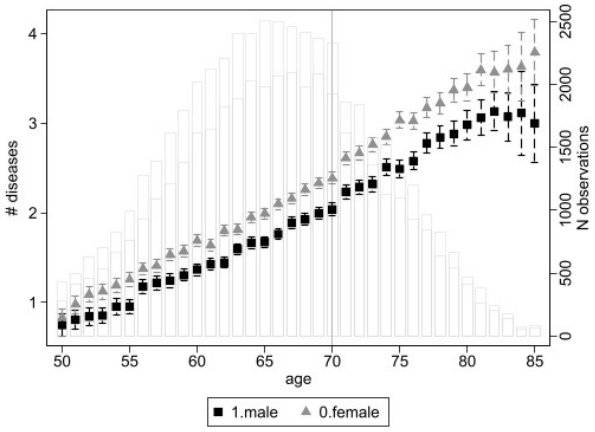
```
diff_d_cancr   85364      2  .0233002    0    1  1st diff of d_cancr ('ever had' | medication)
diff_d_strok   85371      2  .0167856    0    1  1st diff of d_strok ('ever had' | medication)
diff_d_arthr   85430      2  .0697881    0    1  1st diff of d_arthr ('ever had' | medication)
diff_hibper    85467      2  .0667392    0    1  1st diff of hibper ('ever had' - raw data)
diff_diaber    85407      2  .0273865    0    1  1st diff of diaber ('ever had' - raw data)
diff_hearter   85394      2  .0378949    0    1  1st diff of hearter ('ever had' - raw data)
diff_lunger    85371      2  .0207916    0    1  1st diff of lunger ('ever had' - raw data)
diff_psycher   64556      3  .0003563   -1    1  1st diff of psycher ('ever had' - raw data)
diff_osteoer   85355      2  .0087751    0    1  1st diff of osteoer ('ever had' - raw data)
diff_cancrer   85364      2  .0233002    0    1  1st diff of cancrer ('ever had' - raw data)
diff_stroker   85371      2  .0167856    0    1  1st diff of stroker ('ever had' - raw data)
diff_arthrer   85430      2  .0697881    0    1  1st diff of arthrer ('ever had' - raw data)
diff_miss_~p  111212      3  .0775006   -1    1  1st diff of d_hibp ('ever had' | medication) (adj. for gaps)
diff_mis~per  111056      2  .0764119    0    1  1st diff of hibp (ever had - raw data) (adj. for gaps)
diff_miss_~b  111194      3  .0328705   -1    1  1st diff of d_diab ('ever had' | medication) (adj. for gaps)
diff_mis~ber  110986      2  .0320221    0    1  1st diff of diab (ever had - raw data) (adj. for gaps)
diff_miss~rt  111188      3  .0550869   -1    1  1st diff of d_heart ('ever had' | medication) (adj. for gaps)
diff_mis~ter  110967      2  .0405346    0    1  1st diff of heart (ever had - raw data) (adj. for gaps)
diff_miss_~g  111180      3  .0250135   -1    1  1st diff of d_lung ('ever had' | medication) (adj. for gaps)
diff_mis~ger  110943      2  .0242377    0    1  1st diff of lung (ever had - raw data) (adj. for gaps)
diff_miss_~h  111299      3  .0148878   -1    1  1st diff of d_psych ('ever had' | medication) (adj. for gaps)
diff_mis~her   93648      3  .0053391   -1    1  1st diff of psych (ever had - raw data) (adj. for gaps)
diff_miss_~o  111176      3  .0243488   -1    1  1st diff of d_osteo ('ever had' | medication) (adj. for gaps)
diff_mis~oer  110926      2  .0092494    0    1  1st diff of osteo (ever had - raw data) (adj. for gaps)
diff_miss~cr  110934      2  .0264752    0    1  1st diff of d_cancr ('ever had' | medication) (adj. for gaps)
diff_mi~crer  110934      2  .0264752    0    1  1st diff of cancr (ever had - raw data) (adj. for gaps)
diff_miss_~k  110940      2  .0185686    0    1  1st diff of d_strok ('ever had' | medication) (adj. for gaps)
diff_mis~ker  110940      2  .0185686    0    1  1st diff of strok (ever had - raw data) (adj. for gaps)
diff_miss~hr  111016      2  .0769709    0    1  1st diff of d_arthr ('ever had' | medication) (adj. for gaps)
diff_mi~hrer  111016      2  .0769709    0    1  1st diff of arthr (ever had - raw data) (adj. for gaps)
-------------------------------------------------------------------------------------------------------------------------------------------------------

. sum d_anyatfirstobs if sfull & wave==inw_first & agemin==50  // ppl w/ >1 conditions at baseline

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
d_anyatfir~s |        713    .4950912    .5003269          0          1

. sum d_anyatfirstobs if sfull & wave==inw_first & inrange(agemin,50,65) // ppl w/ >1 conditions at baseline

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
d_anyatfir~s |     14,087    .6508128    .4767301          0          1

. qui log close log
```

4 Figures and Tables

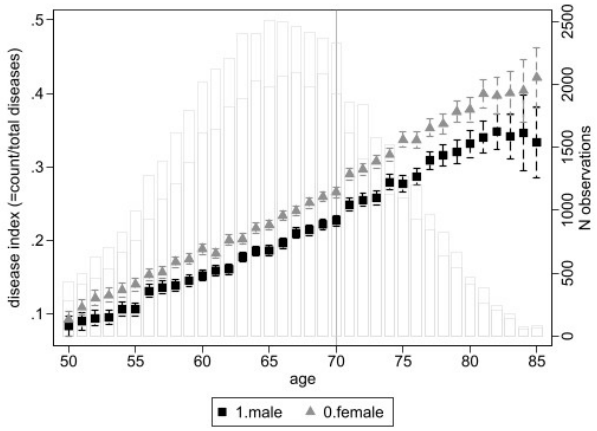
Figure 1: Variables by age (pooled sample) with 95% CI



(a) d_any (Any Disease)



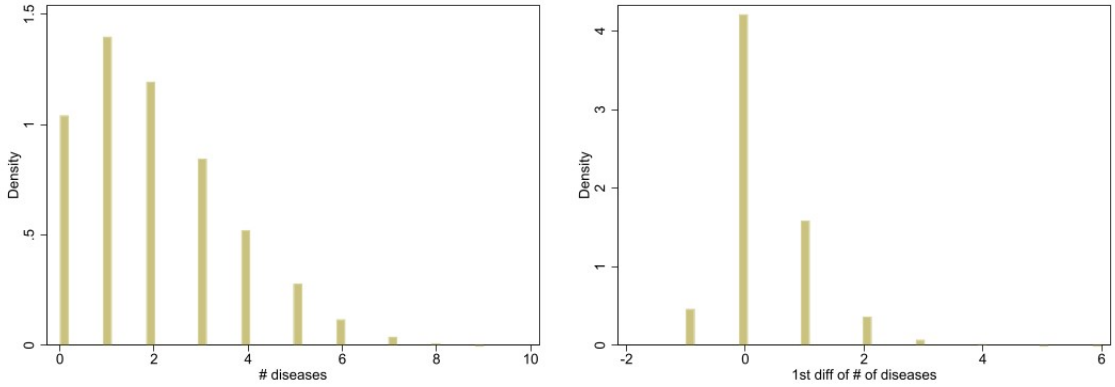
(b) d_count (number of diseases)



(c) d_count_index (count / considered diseases)

Note: ...

Figure 2: Distributions of variables

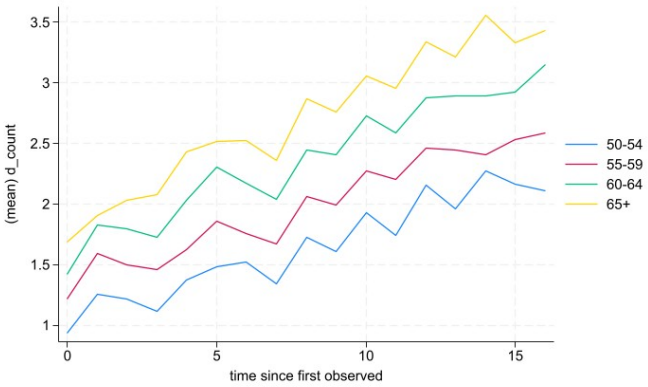


(a) d_count

(b) diff_d_count (1st difference in count)

```
------------------------------------------------------------------------------------------------------------------------------------------------
      name:  log
       log:  C:/Users/User/Documents/GitHub/2-projectMM-SHARE/files/logs/log-g_bytime-cohortmin5.txt
  log type:  text
 opened on:  21 Jan 2024, 20:46:06

. ** xtline by age group **
.
. loc      sample "sfull"

. loc      timevar "timesincefirstobs_yr" // timesincefirstobs_yr | time

. collapse (mean) d_count_mean = d_count if `sample'==1, by(cohortmin5 `timevar')

. xtset    cohortmin5 `timevar'

Panel variable: cohortmin5 (unbalanced)
 Time variable: timesincefirstobs_yr, 0 to 15
         Delta: 1 unit

. loc       y     "d_count_mean"

. xtline   `y', overlay

. gr export       "$outpath/fig/supplement/g_by`timevar'-cohortmin5_`sample'_d_count.jpg", replace
(file C:/Users/User/Documents/GitHub/2-projectMM-SHARE/files/fig/supplement/g_bytimesincefirstobs_yr-cohortmin5_sfull_d_count.
> jpg not found)
file C:/Users/User/Documents/GitHub/2-projectMM-SHARE/files/fig/supplement/g_bytimesincefirstobs_yr-cohortmin5_sfull_d_count.j
> pg written in JPEG format

. qui log close log
```

Figure 3: Count of Diseases by age at baseline



(a) timevar: timesincefirstobs

4.1 Ordered Response Models

```
------------------------------------------------------------------------------------------------------------------------------------
      name:  log
       log:  C:/Users/User/Documents/GitHub/2-projectMM-SHARE/files/logs/log-t-regd_count-cohort.txt
  log type:  text
 opened on:  21 Jan 2024, 23:39:02

.         tab d_count wave if sfull

           |                                Survey Wave
# diseases | 2004 wave  2006/07 w  2011/12 w  2013 wave  2015 wave  2017 wave  2019/20 w |     Total
-----------+--------------------------------------------------------------------------------+----------
         0 |     3,584      3,906      3,277      2,653      1,994      1,677      1,002 |    18,093
         1 |     3,289      4,100      4,310      3,783      3,331      2,923      2,000 |    23,736
         2 |     1,857      2,750      3,346      3,408      3,355      3,097      2,255 |    20,068
         3 |       933      1,549      2,106      2,297      2,582      2,629      2,019 |    14,115
         4 |       393        782      1,149      1,414      1,622      1,780      1,496 |     8,636
         5 |       126        330        530        752        954        977        918 |     4,587
         6 |        41        105        173        286        419        428        470 |     1,922
         7 |         8         25         37         86        145        165        167 |       633
         8 |         3          7          7         14         30         43         43 |       147
         9 |         0          0          1          2          3          3          3 |        12
-----------+--------------------------------------------------------------------------------+----------
     Total |    10,234     13,554     14,936     14,695     14,435     13,722     10,373 |    91,949

.         *** ordered logit (wave 1 only)***
.         eststo m1: ologit `y' c.age##c.age `ctrls'              if `sample'==1 & wave==1 , vce(robust)

Iteration 0:  Log pseudolikelihood = -15011.882
Iteration 1:  Log pseudolikelihood = -14572.135
Iteration 2:  Log pseudolikelihood = -14570.006
Iteration 3:  Log pseudolikelihood = -14570.005

Ordered logistic regression                             Number of obs = 10,230
                                                        Wald chi2(6)  = 869.73
                                                        Prob > chi2   = 0.0000
Log pseudolikelihood = -14570.005                       Pseudo R2     = 0.0294

------------------------------------------------------------------------------------------
                           |               Robust
                   d_count | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
---------------------------+--------------------------------------------------------------
                       age |   .1562847   .0500177     3.12   0.002     .0582517    .2543176
                           |
                 c.age#c.age |  -.0007876   .0004219    -1.87   0.062    -.0016145    .0000393
                           |
                      male |  -.4183029   .0371905   -11.25   0.000    -.4911949   -.3454109
                  marriedr |  -.2550336    .046066    -5.54   0.000    -.3453212   -.1647459
                           |
                   raeducl |
2.upper secondary or vocational  |  -.3350006   .0421869    -7.94   0.000    -.4176854   -.2523158
      3.tertiary education  |  -.5680077   .0474354   -11.97   0.000    -.6609794    -.475036
---------------------------+--------------------------------------------------------------
                     /cut1 |   5.173221   1.470574                      2.290949    8.055492
                     /cut2 |   6.596659   1.470752                      3.714037    9.479281
                     /cut3 |   7.694131   1.470425                      4.812151    10.57611
                     /cut4 |    8.79711   1.470895                      5.914208    11.68001
                     /cut5 |    10.0289    1.47079                        7.1462    12.91159
                     /cut6 |   11.31966    1.47675                      8.425278    14.21403
                     /cut7 |   13.13849   1.516574                      10.16606    16.11092
------------------------------------------------------------------------------------------

. //  margins, predict(outcome(5))
.         margins, // at(mpg = (10(10)40)) , dydx(male)

Predictive margins                                      Number of obs = 10,230
Model VCE: Robust

1._predict: Pr(d_count==0), predict(pr outcome(0))
2._predict: Pr(d_count==1), predict(pr outcome(1))
3._predict: Pr(d_count==2), predict(pr outcome(2))
4._predict: Pr(d_count==3), predict(pr outcome(3))
5._predict: Pr(d_count==4), predict(pr outcome(4))
6._predict: Pr(d_count==5), predict(pr outcome(5))
7._predict: Pr(d_count==6), predict(pr outcome(6))
```

```
8._predict: Pr(d_count==7), predict(pr outcome(7))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    _predict |
           1 |   .3490889   .0045661    76.45   0.000     .3401396    .3580383
           2 |   .3209133   .0046117    69.59   0.000     .3118746    .3299521
           3 |    .182286   .0037785    48.24   0.000     .1748804    .1896917
           4 |   .0918929   .0028364    32.40   0.000     .0863337     .097452
           5 |   .0386652   .0019011    20.34   0.000     .0349391    .0423912
           6 |   .0123579    .001092    11.32   0.000     .0102175    .0144983
           7 |   .0040137   .0006251     6.42   0.000     .0027886    .0052389
           8 |    .000782   .0002764     2.83   0.005     .0002403    .0013238
------------------------------------------------------------------------------

. //      marginsplot
. //      predict p0ologit p1ologit p2ologit p3ologit p4ologit p5ologit p6ologit p7ologit p8ologit, pr // p9
. //      sum     p?ologit
.
. //      *findit spost13 // needed for -mtable-
.         *mtable, at(male = (0 1))
.
.         **brant test**
. //      *https://www.statalist.org/forums/forum/general-stata-discussion/general/1335252-ologit-and-brant-test
.         brant

Brant test of parallel regression assumption

             |       chi2     p>chi2      df
-------------+------------------------------
         All |      63.92      0.003      36
-------------+------------------------------
         age |      12.74      0.047       6
  c.age#c.age |      11.91      0.064       6
        male |       2.18      0.903       6
    marriedr |       8.03      0.236       6
   2.raeducl |      15.64      0.016       6
   3.raeducl |      11.52      0.074       6

A significant test statistic provides evidence that the parallel
regression assumption has been violated.

.
.         ** xt-ordered logit **
.         eststo m2: xtologit ‘y’ c.age##c.age ‘ctrls’    if ‘sample’==1,  vce(cluster ID) // -vce(cl ID)- is equivalent to -r
> obust-

Fitting comparison model:

Iteration 0:  Log likelihood = -162992.84
Iteration 1:  Log likelihood = -154909.45
Iteration 2:  Log likelihood = -154814.28
Iteration 3:  Log likelihood = -154814.17
Iteration 4:  Log likelihood = -154814.17

Refining starting values:

Grid node 0:  Log likelihood =  -139029.5

Fitting full model:

Iteration 0:  Log pseudolikelihood =  -139029.5
Iteration 1:  Log pseudolikelihood = -121713.65
Iteration 2:  Log pseudolikelihood = -117069.73
Iteration 3:  Log pseudolikelihood = -116329.31
Iteration 4:  Log pseudolikelihood = -116248.74
Iteration 5:  Log pseudolikelihood =  -116247.6
Iteration 6:  Log pseudolikelihood =  -116247.6

Random-effects ordered logistic regression        Number of obs    =   91,698
Group variable: ID                                Number of groups =   18,852

Random effects u_i ~ Gaussian                     Obs per group:
                                                              min =        1
```

```
                                                      avg =      4.9
                                                      max =        7

Integration method: mvaghermite                 Integration pts. =       12

                                                Wald chi2(6)     = 17301.97
Log pseudolikelihood = -116247.6                Prob > chi2      =   0.0000

                                                  (Std. err. adjusted for 18,852 clusters in ID)
------------------------------------------------------------------------------------------------
                               |               Robust
                     d_count | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------------------+----------------------------------------------------------------
                           age |   .1069406   .0252989     4.23   0.000     .0573556    .1565256
                               |
                     c.age#c.age |   .0015939   .0001914     8.33   0.000     .0012187    .0019691
                               |
                          male |  -.7894021   .0604117   -13.07   0.000    -.9078068   -.6709974
                       marriedr |  -.3674242   .0508499    -7.23   0.000    -.4670883   -.2677602
                               |
                        raeducl |
 2.upper secondary or vocational  |  -.5155526   .0676206    -7.62   0.000    -.6480865   -.3830187
            3.tertiary education  |  -1.446978   .0792749   -18.25   0.000    -1.602354   -1.291602
-------------------------------+----------------------------------------------------------------
                         /cut1 |   8.759927   .8302374                      7.132691    10.38716
                         /cut2 |   12.11553   .8325494                      10.48376    13.74729
                         /cut3 |   14.65638   .8336741                      13.02241    16.29035
                         /cut4 |   16.89966    .833902                      15.26525    18.53408
                         /cut5 |   19.03248   .8336916                      17.39847    20.66648
                         /cut6 |   21.22754   .8342634                      19.59241    22.86267
                         /cut7 |   23.60316   .8359843                      21.96466    25.24166
-------------------------------+----------------------------------------------------------------
                      /sigma2_u |   15.16607   .2375009                      14.70765    15.63878
------------------------------------------------------------------------------------------------

. *       brant // does not work with xtologit
.
.         ** gologit2 **
.         eststo m3: qui gologit2 ‘y’ c.age##c.age male           if ‘sample’==1 & wave==1, vce(cluster ID) autofit  // gologi
> t without controls
--Break--
r(1);

end of do-file

--Break--
r(1);

. do "C:\Users\User\AppData\Local\Temp\STD136c_000000.tmp"

.         qui log close log
```

5 THE PART BELOW CONTAINS PRELIMINARY RESULTS. IT IS SYNCED TO GITHUB AND CAN BE DISREGARDED UNTIL ADDED ABOVE. — For the team to edit/add notes to the part above, use "7-graphsAndTables.tex"