Turkey 2008.
Note: Yadigar changed some cases from de jure to de facto (or something like that - see "wealth index problem.doc".
As a result, need to re-run wealth index to make sure the distribution returns to 20% in each quintile.
Re-running on 6-11-09, and coming back to the issues on 6-22-09.

GET
FILE='C:\Users\Sarah.E.Bradley\Documents\Other Data\Wealth index\Turkey 2\'008\TKHR52-6-23-09.sav'.
DATASET NAME DataSet3 WINDOW=FRONT.

FREQ hv015.
* select only completed interviews.
SELECT IF hv015 = 1.
freq hv015.

*Note that the following variables have been changed in the new dataset:
* SH142B - # sleeping rooms ->HV216.
* SH144A - Refrigerator -> HV209.
* SH144L - Television -> HV208.
* SH144S - Telephone (non-mob) -> HV221.
* SH144Z2 - Motorcycle/scooter -> HV211.

**New version of freqs, Deleting any variables that are missing or empty.
Freq  HV204 HV205  HV208 HV209 HV211 HV212
      HV213 HV216 HV221 HV225
      HV242 HV243A SH123 SH125 sh129
      SH137 SH138 SH141 SH142A SH142D SH143 SH144B SH144C SH144D
      SH144E SH144F SH144G SH144H SH144I SH144J SH144K SH144M SH144N
      SH144O SH144P SH144Q SH144R SH144T SH144U SH144V SH144W SH144X SH144Y
      SH144Z SH144Z1 .

***# of HH MEMBERS (for sleeping rooms and HHmemweight).
*Check to see if any HHs have 0 de jure members, and if so sub in de facto members.
freq hv012 hv013.
compute nummem=hv012.
if (hv012=0) nummem=hv013.
freq nummem.

**HH MEMBERS over age 15 (for cellphones).
compute numadult15=0.
if (adult>=15 & adult<97) numadult15=numadult15+1.
end repeat.
freq numadult15.

compute numadult152=0.
do repeat adult=HV105$01 to HV105$35.
if (adult>=15 & adult<97) numadult152=numadult152+1.
end repeat.
freq numadult152.

**HH MEMBERS over age 18 (for cars). 
compute numadult18=0.
if (adult>=18 & adult<97) numadult18=numadult18+1.
end repeat.
freq numadult18.
**NOTE: there are 5 HHs with no adults 18+; assign them a value
of 1.
if (numadult18=0) numadult18=1.
freq numadult18.

* WATER.
freq sh129.
COMPUTE h2opipe = 0.
IF (sh129 = 11) h2opipe = 1.
VAR LABELS h2opipe "if gets water piped into home".
VAL LABELS h2opipe 0 "no water piped into home"
1 "water is piped into home".
CROSSTABS TABLES=h2opipe BY SH129.

COMPUTE h2opubtap = 0.
IF (sh129=12) h2opubtap = 1.
VAR LABELS h2opubtap "if gets water thru public tap".
VAL LABELS h2opubtap 0 "no water public tap" 1 "water is public tap".

COMPUTE h2ppvwel = 0.
IF (sh129 = 21) h2ppvwel = 1.
VAR LABELS h2ppvwel "if gets water from a protected well-in residence".
VAL LABELS h2ppvwel 0 "no water from a protected well" 1 "water is from a protected well".

COMPUTE h2pydwel = 0.
IF (sh129 = 22) h2pydwel = 1.
VAR LABELS h2pydwel "if gets water from an unprotected well-publicl".
VAL LABELS h2pydwel 0 "no water from an unprotected well" 1 "water is from an unprotected well".

COMPUTE h2osurf = 0.
IF (sh129 = 31) h2osurf = 1.
VAR LABELS h2osurf "if gets water from a surface source".
VAL LABELS h2osurf 0 "no water from a surface source" 1 "water is from a surface source".

COMPUTE h2spring = 0.
IF (sh129=32) h2spring = 1.
VAR LABELS h2spring "if gets water from a spring or pub fountain".
VAL LABELS h2spring 0 "no water from a spring" 1 "water is from a spring".

COMPUTE h2riv = 0.
IF (sh129=33) h2riv = 1.
VAR LABELS h2riv "if gets water from a river/stream/pond/lake".
VAL LABELS h2riv 0 "no water from a river" 1 "water is from a river".

COMPUTE h2otruck = 0.
IF (sh129=51) h2otruck = 1.
VAR LABELS h2otruck "if gets water from truck".
VAL LABELS h2otruck 0 "no water from truck" 1 "water is from truck".

COMPUTE h2obot = 0.
IF (sh129 = 61) h2obot = 1.
VAR LABELS h2obot "if gets water from bottle".
VAL LABELS h2obot 0 "no water from bottle" 1 "water is from bottle".

COMPUTE h2ooth = 0.
IF (sh129=96 | sh129=99) h2ooth = 1.
VAR LABELS h2ooth "if gets water from other".
VAL LABELS h2ooth 0 "no water from other"
1 "water is from other".
*IF (MISSING(sh129)) h2ooth = 1.
*note: removed missing from combination with other due to non pos def matrix.
crosstab tables=h2ooth by sh129.
freq sh129 h2ooth.
*excluded non-drinking water due to too few cases in many categories.
*TOILET.
freq sh137 sh138 hv225.
crosstab tables=sh138 by hv225.
crosstab tables=sh138 by sh137.
crosstab tables=sh137 by hv225 by sh138.
compute tbush=0.
if (sh137=0) tbush=1.
var labels tbush "uses bush or public toilet".
val labels tbush 0 "does not use bush or pub toilet"
1 "uses bush or pub toilet".
crosstab tables sh137 by tbush.
compute tins=0.
if (sh137=1) tins=1.
var labels tins "uses inside toilet".
val labels tins 0 "does not use inside toilet"
1 "uses inside toilet".
crosstab tables sh137 by tins.
compute tout=0.
if (sh137=2) tout=1.
var labels tout "uses outside toilet".
val labels tout 0 "does not use outside toilet"
1 "uses outside toilet".
crosstab tables sh137 by tout.
compute tboth=0.
if (sh137=3) tboth=1.
var labels tboth "uses in and outside toilet".
val labels tboth 0 "does not use in and outside toilet"
1 "uses inside and outside toilet".
crosstab tables sh137 by tboth.
COMPUTE flpvts = 0.
IF (sh138=1 & hv225 = 0) flpvts = 1.
VAR LABELS flpvts "if uses pvt flush toilet to sewer".
VAL LABELS flpvts 0 "does not use pvt flush toilet"
1 "uses pvt flush toilet".
crosstab tables sh138 by hv225 by flpvts.

COMPUTE flshrs = 0.
IF (sh138=1 & hv225 = 1) flshrs = 1.
VAR LABELS flshrs "if uses shared flush toilet to sewer".
VAL LABELS flshrs 0 "does not use shared flush toilet"
1 "uses shared flush toilet".
crosstab tables sh138 by hv225 by flshrs.

COMPUTE opitpvt = 0.
IF (sh138=2 & hv225 = 0) opitpvt = 1.
VAR LABELS opitpvt "if uses pvt open pit".
VAL LABELS opitpvt 0 "does not use pvt open pit"
1 "uses pvt open pit".
crosstab tables sh138 by hv225 by opitpvt.

COMPUTE opitpub = 0.
IF (sh138=2 & hv225 = 1) opitpub = 1.
VAR LABELS opitpub "if uses public open pit".
VAL LABELS opitpub 0 "does not use pub open pit"
1 "uses pub open pit".
crosstab tables sh138 by hv225 by opitpub.

COMPUTE cpitpvt = 0.
IF (sh138=3 & hv225 = 0) cpitpvt = 1.
VAR LABELS cpitpvt "if uses pvt closed pit".
VAL LABELS cpitpvt 0 "does not use pvt closed pit"
1 "uses pvt closed pit".
crosstab tables sh138 by hv225 by cpitpvt.

COMPUTE cpitpub = 0.
IF (sh138=3 & hv225 = 1) cpitpub = 1.
VAR LABELS cpitpub "if uses public closed pit".
VAL LABELS cpitpub 0 "does not use pub closed pit"
1 "uses pub closed pit".
crosstab tables sh138 by hv225 by cpitpub.

compute tother=0.
if (sh138=7) tother=1.
var labels tother "uses other toilet type-private +1 public".
val labels tother 0 "does not use other toilet"
1 "uses other toilet".
crosstab tables sh138 by hv225 by tother.

*HEAT.
freq sh141.

compute hcengas=0.
if (sh141=11 | sh141=12) hcengas=1.
var labels hcengas "heat is central natural gas + 9 diesel gas".
val labels hcengas 0 "does not use central gas heat"
1 "uses central gas heat".
crosstab tables sh141 by hcengas.

compute hcenwood=0.
if (sh141=13) hcenwood=1.
var lab hcenwood "heat is central wood or coal".
val lab hcenwood 0 "no central wood heat"
 1 "uses central wood heat".
crosstab tables sh141 by hcenwood.

compute hcenoth=0.
if (sh141=14) hcenoth=1.
var lab hcenoth "heat is central other".
val lab hcenoth 0 "no central other heat"
 1 "uses central other heat".
crosstab tables sh141 by hcenoth.

compute hflngas=0.
if (sh141=21) hflngas=1.
var lab hflngas "heat is flat natural gas".
val lab hflngas 0 "no flat natural gas heat"
 1 "uses flat natural gas heat".
crosstab tables sh141 by hflngas.

compute hflnoil=0.
if (sh141=22) hflnoil=1.
var lab hflnoil "heat is flat diesel oil".
val lab hflnoil 0 "no flat diesel oil heat"
 1 "uses flat diesel oil heat".
crosstab tables sh141 by hflnoil.

compute hstgas=0.
if (sh141=31 | sh141=32) hstgas=1.
var lab hstgas "heat is stove natural gas + 12 stove diesel".
val lab hstgas 0 "no stove natural gas heat"
 1 "uses stove natural gas heat".
crosstab tables sh141 by hstgas.

compute hstwood=0.
if (sh141=33) hstwood=1.
var lab hstwood "heat is stove wood or coal".
val lab hstwood 0 "no stove wood heat"
 1 "uses stove wood heat".
crosstab tables sh141 by hstwood.

compute hsttdung=0.
if (sh141=34) hsttdung=1.
var lab hsttdung "heat is stove dried cow dung".
val lab hsttdung 0 "no stove dung heat"
 1 "uses stove with dung for heat".
crosstab tables sh141 by hsttdung.
compute hstother=0.
if (sh141=35) hstother=1.
var lab hstother "heat is stove other fuel".
val lab hstother 0 "no stove other"
  1 "uses stove other for heat".
crosstab tables sh141 by hstother.

compute helec=0.
if (sh141=40) helec=1.
var lab helec "heat is electric".
val lab helec 0 "no electric heat"
  1 "uses electric heat".
crosstab tables sh141 by helec.

compute hother=0.
if (sh141=96) hother =1.
var lab hother "heat is other ".
val lab hother 0 "no other heat"
  1 "uses other heat".
freq hother sh141.
*note: not assigning 10 missing cases.

* FLOORING.
freq sh143.
COMPUTE dirtfloo = 0.
IF (sh143= 11) dirtfloo = 1.
VAR LABELS dirtfloo "if floors are made of earth or sand".
VAL LABELS dirtfloo 0 "floors are not made of earth"
  1 "floors are made of earth".
crosstab tables sh143 by dirtfloo.

COMPUTE woodfloo = 0.
IF (sh143 = 21) woodfloo = 1.
VAR LABELS woodfloo "if floors are made of wood planks".
VAL LABELS woodfloo 0 "floors are not made of wood planks"
  1 "floors are made of wood planks".
crosstab tables sh143 by woodfloo.

COMPUTE parqfloo = 0.
IF (sh143 = 31) parqfloo = 1.
VAR LABELS parqfloo "if floors are made of parquet, polished wood".
VAL LABELS parqfloo 0 "floors are not made of parquet, polished wood"
  1 "floors are made of parquet, polished wood".
crosstab tables sh143 by parqfloo.

COMPUTE karofloo = 0.
IF (sh143=32) karofloo = 1.
VAR LABELS karofloo "if floors are made of Karo".
VAL LABELS karofloo 0 "floors are not made of Karo"
  1 "floors are made of Karo".
crosstab tables sh143 by karofloo.

COMPUTE cemtfloo = 0.
IF (sh143=34) cemtfloo = 1.
VAR LABELS cemtfloo "if floors are made of cement ".
VAL LABELS cemtfloo 0 "floors are not made of cement"
   1 "floors are made of cement".
crosstab tables sh143 by cemtfloo.

COMPUTE carpetfloo = 0.
IF (sh143=35) carpetfloo = 1.
VAR LABELS carpetfloo "if floors are made of carpet ".
VAL LABELS carpetfloo 0 "floors are not made of carpet"
   1 "floors are made of carpet".
crosstab tables sh143 by carpetfloo.

COMPUTE marleyfloo = 0.
IF (sh143=36) marleyfloo = 1.
VAR LABELS marleyfloo "if floors are made of marley".
VAL LABELS marleyfloo 0 "floors are not made of marley"
   1 "floors are made of marley".
crosstab tables sh143 by marleyfloo.

COMPUTE mozfloo = 0.
IF (sh143=37) mozfloo = 1.
VAR LABELS mozfloo "if floors are made of mozaic".
VAL LABELS mozfloo 0 "floors are not made of mozaic"
   1 "floors are made of mozaic".
crosstab tables sh143 by mozfloo.

COMPUTE laminatefloo = 0.
IF (sh143=38) laminatefloo = 1.
VAR LABELS laminatefloo "if floors are made of laminate".
VAL LABELS laminatefloo 0 "floors are not made of laminate"
   1 "floors are made of laminate".
crosstab tables sh143 by laminatefloo.

COMPUTE otherfloo = 0.
IF (sh143=96) otherfloo = 1.
VAR LABELS otherfloo "if floors are made of other".
VAL LABELS otherfloo 0 "floors are not made of other"
   1 "floors are made of other".
*IF (MISSING(sh143)) otherfloo = 1.
crosstab tables sh143 by otherfloo.
freq otherfloo sh143.

* If kitchen is separate room.
COMPUTE kitchen = 0.
IF (hv242 = 1) kitchen = 1.
VAR LABELS kitchen "if has separate room for kitchen in hh".

freq kitchen sh143.
VAL LABELS kitchen 0 "no separate kitchen"
    1 "has separate room for kitchen".
crosstab tables hv242 by kitchen.

* If bathroom is separate room.
freq sh142d.
COMPUTE bathroom = 0.
IF (sh142D) bathroom = 1.
VAR LABELS bathroom "if has separate room for bathroom in hh".
VAL LABELS bathroom 0 "no separate bathroom"
    1 "has separate room for bathroom".
crosstab tables sh142d by bathroom.

*ROOMS PER HH MEMBER.
*NOTE: updating from sh142b to hv216.
freq sh142a hv216.
*recoding 1 obs of 64 rooms to missing - huge outlier.
recode sh142a (64=sysmis) (missing=sysmis).
recode hv216 (99=sysmis).

freq hv012.
COMPUTE memsleep = (nummem/hv216).
IF (MISSING(hv216)) memsleep = hv012.
VARIABLE LABELS memsleep "number of members per sleeping room".
freq memsleep.

COMPUTE memroom = (nummem/sh142a).
IF (MISSING(sh142a)) memroom = hv012.
VARIABLE LABELS memroom "number of members per total room".
freq memroom.
EXECUTE.

*number of cell phones and cars.
freq sh144r sh144y.
compute numcellphone=sh144r.
compute numcar=sh144y.
exe.

**# of people>=15/cell phone.
compute cellphonemem=(numcellphone/numadult15).
freq cellphonemem.
means cellphonemem by numadult15 /cells=mean.

**# people>=18/car.
compute carmem=(numcar/numadult18).
freq carmem.
means carmem by numadult18 /cells=mean.
crosstab tables carmem by numcar.

*has TV.
freq hv208.
compute TV=hv208.

*Refrigerator.
freq hv209.
compute refrigerator=hv209.

*motorcycle/scooter.
freq hv211.
compute motorcycle=hv211.

*land line telephone.
freq hv221.
compute telephone=hv221.

*gas/elec oven.
freq sh144b.
compute oven=sh144b.

*microwave.
freq sh144c.
compute microwave=sh144c.

*blender.
freq sh144d.
compute blender=sh144d.

*dishwasher.
freq sh144e.
compute dishwasher=sh144e.

*garbage grinder.
freq sh144f.
compute grinder=sh144f.

*washer.
freq sh144g.
compute washer=sh144g.

*dryer.
freq sh144h.
compute dryer=sh144h.

*iron.
freq sh144i.
compute iron=sh144i.

*vacuum.
freq sh144j.
compute vacuum=sh144j.

*LCD/plasma tv.
freq sh144k.
compute plasma=sh144k.

cable.
freq sh144m.
compute cable=sh144m.

*satellite.
freq sh144n.
compute satellite=sh144n.

*video camera.
freq sh144o.
compute video=sh144o.

*DVD player.
freq sh144p.
compute DVD=sh144p.

camera.
freq sh144q.
compute camera=sh144q.

*laptop.
freq sh144t.
compute laptop=sh144t.

*are laptop and computer same?.
crosstab tables sh144t by sh144u.
*NO.

*computer.
freq sh144u.
compute computer=sh144u.

*internet.
freq sh144v.
compute internet=sh144v.

*Fitness equipment.
freq sh144w.
compute fitness=sh144w.

*Air conditioner.
freq sh144x.
compute AC=sh144x.

*Taxi.
freq sh144z.
compute taxi=sh144z.

*Tractor.
freq sh144z1.
compute tractor=sh144z1.

*Simple assets:.
FREQUENCIES
  VARIABLES=kitchen bathroom numcellphone numcar TV refrigerator motorcycle telephone oven blender dishwasher grinder washer dryer vacuum plasma cable satellite video DVD camera laptop computer internet fitness AC taxi microwave iron tractor
  /ORDER= ANALYSIS .

*replace missing w don't have:.
IF (MISSING(kitchen)) kitchen = 0.
IF (MISSING(bathroom)) bathroom = 0.
IF (MISSING(numcellphone)) numcellphone = 0.
IF (MISSING(numcar)) numcar  = 0.
IF (MISSING(TV)) TV= 0.
IF (MISSING(refrigerator)) refrigerator = 0.
IF (MISSING(motorcycle)) motorcycle = 0.
IF (MISSING(telephone )) telephone= 0.
IF (MISSING(oven)) oven  = 0.
IF (MISSING(blender )) blender  = 0.
IF (MISSING(dishwasher)) dishwasher= 0.
IF (MISSING(grinder)) grinder = 0.
IF (MISSING(washer)) washer = 0.
IF (MISSING(dryer)) dryer = 0.
IF (MISSING(vacuum)) vacuum = 0.
IF (MISSING(plasma)) plasma = 0.
IF (MISSING(cable)) cable = 0.
IF (MISSING(satellite)) satellite = 0.
IF (MISSING(video)) video= 0.
IF (MISSING(DVD )) DVD = 0.
IF (MISSING(camera)) camera = 0.
IF (MISSING(laptop)) laptop= 0.
IF (MISSING(computer)) computer = 0.
IF (MISSING(internet)) internet = 0.
IF (MISSING(fitness)) fitness = 0.
IF (MISSING(AC)) AC = 0.
IF (MISSING(taxi)) taxi = 0.
IF (MISSING(microwave)) microwave = 0.
IF (MISSING(iron)) iron = 0.
IF (MISSING(tractor)) tractor = 0.

*Simple assets:.
FREQUENCIES
VARIABLES=kitchen bathroom numcellphone numcar TV refrigerator motorcycle telephone oven blender dishwasher grinder washer dryer vacuum plasma cable satellite video DVD camera laptop computer internet fitness AC taxi microwave iron tractor
/ORDER= ANALYSIS.

*complex assets:.
FREQUENCIES
VARIABLES=h2opipe h2opubtap h2ppvwel h2pydwel h2osurf h2spring h2riv
h2otruck h2obot h2ooth tbush tins tout tboth flpvts flshrs opitpvt opitpub
 cpitpvt cpitpub tother hcengas hcenwood hcenoth hflngas hflnoil hstgas
 hstwood hsttdung hstother helec hother dirtflool woodfloo parqfloo karofloo
cemtfloo carpetfloo marleyfloo mozfloo laminatefloo otherfloo memsleep memroom
/ORDER= ANALYSIS.

*NOTE: excluded tins tout tboth for first run, then crosstab - they are correlated and improve eigenvalue, so included.
**NOTE: tested removing internet, satellite and cable to see if that made the wealth distribution more even over urban-rural
   * Instead, made the issue worse (correlation increased and % of rural in highest quintile decreased).

FACTOR
/VARIABLES h2opipe h2opubtap h2ppvwel h2pydwel h2osurf h2spring h2riv
h2otruck h2obot h2ooth tbush flpvts flshrs opitpvt opitpub
 cpitpvt cpitpub tother hcengas hcenwood hcenoth hflngas hflnoil hstgas
 hstwood hsttdung hstother helec hother dirtflool woodfloo parqfloo karofloo
cemtfloo carpetfloo marleyfloo mozfloo laminatefloo otherfloo memsleep memroom
 kitchenn bathroom cellphonemem carmem TV refrigerator motorcycle telephone oven blender dishwasher grinder washer dryer vacuum plasma cable satellite video DVD camera laptop computer internet fitness AC taxi
 microwave iron tractor tins tout tboth
/MISSING MEANSUB /ANALYSIS h2opipe h2opubtap h2ppvwel h2pydwel h2osurf h2spring h2riv
h2otruck h2obot h2ooth tbush flpvts flshrs opitpvt opitpub
 cpitpvt cpitpub tother hcengas hcenwood hcenoth hflngas hflnoil
save outfile="TKassets6-22-09.sav".
COMPUTE hhmemwt = hv005/1000000 * nummem .
VARIABLE LABELS hhmemwt 'HH members weighting for Index' .

WEIGHT
BY hhmemwt .
FREQUENCIES
VARIABLES=fac1_1 /FORMAT=NOTABLE
/NTILES= 5
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN /ORDER ANALYSIS .
freq hv012.

RANK
   VARIABLES=FAC1_1 (A)
   /RANK /NTILES (5)
   /PRINT=YES
   /TIES=MEAN .

freq nfac1_1.

crosstab tables nfac1_1 by hv025 /cells=row count
/statistics=chisq.

compute wthind5=nfac1_1 .

write outfile="TKscores6-22-09.dat" records=1 table
/hhid fac1_1 wthind5.
execute.

MEANS
   TABLES=h2opipe h2opubtap h2ppvwel h2pydwel h2osurf h2spring h2riv
   h2otruck h2obot h2ooth tbush flpvts flshrs opitpvt opitpub
cpitpvt cpitpub tother hcengas hcenwood hcenoth hflngas hflnoil hstgas
hstwood hsttdung hstother helec hother dirtflool woodflool
parqflool karoflool
cemflool carpetflool marleyflool mozflool laminateflool otherflool
memsleep memroom
kitchen bathroom cellphonemem carmem TV refrigerator motorcycle
telephone oven blender dishwasher grinder washer dryer vacuum
plasma cable
satellite video DVD camera laptop computer internet fitness AC
taxi
microwave iron tractor tins tout tboth
BY
w1thind5
/CELLS MEAN  .

MEANS
  TABLES= hv009 hv025 BY w1thind5
  /CELLS MEAN COUNT STDDEV
  /STATISTICS ANOVA  .

crosstab tables hv025 by w1thind5
  /cell col
  /statistics chisq.

FREQ w1thind5.
weight off.
FREQ w1thind5.