

Part 1. 可視化の導入

1 Data import

Excel 形式のデータのインポートし、この後の解析に向けて簡単な処理を行ないます。

part1.do

```
import excel "NHEFS.xlsx", firstrow clear

///// アウトカム変数へのラベル作成 /////
label define death 0 "Alive" 1 "Death"
label values death death

///// pack_year_n 変数の作成 /////
gen pack_year_n = (smokeintensity/20) * smokeyrs

compress
save nhefs.dta, replace
```

また、必要な外部コマンド (user-driven command または community-based command 等) をインストールして下さい。既にインストール済みのものは省いても差し支えありません。

part1.do

```
ssc install schemepack, replace // イケてるグラフテーマ
ssc install stripplot, replace // dotplot の (上位) 互換コマンド
ssc install fre, replace // 度数分布表。tab の互換コマンド
ssc install table1, replace // 記述統計表を作成
ssc install bihist, replace // 上下に向けたヒストグラム
ssc install covbal, replace // 共変量バランスのチェック
// ssc install dag, replace // DAG を用いた解析 (今回はつかいません)

// グラフを 1 枚にまとめるときに、Legend を 1 つだけにするコマンド grc1leg
net install grc1leg.pkg, replace from(http://www.stata.com/users/vwiggins/)

// labmask のインストール
net install gr0034.pkg, replace from(http://www.stata-journal.com/software/sj8-2/)

// カラーパレット
net install gr0075.pkg, replace from(http://www.stata-journal.com/software/sj18-4/)
```

3 twoway

3.1 twoway について

(主に) 2つの変数の関係を二次元にプロットするためのコマンドです。Rにおける **ggplot** のようにレイヤーを追加して、複数のグラフを1つの座標系にプロット可能です。Stataでは、**//** (スラッシュ2つ) の後にコメントを残す事ができます。また、**///** (スラッシュ3つ) では、その後にコメントを残しつつ、次の行にコマンドが継続することをします。

3.2 基本コマンド

Rにおける **ggplot2** と同じように **twoway** コマンドだけでは、何も描かれていない真っ白い背景だけが用意されます。ここにどんどんとコマンドでグラフに関する指示を重ねていきます。

part1.do

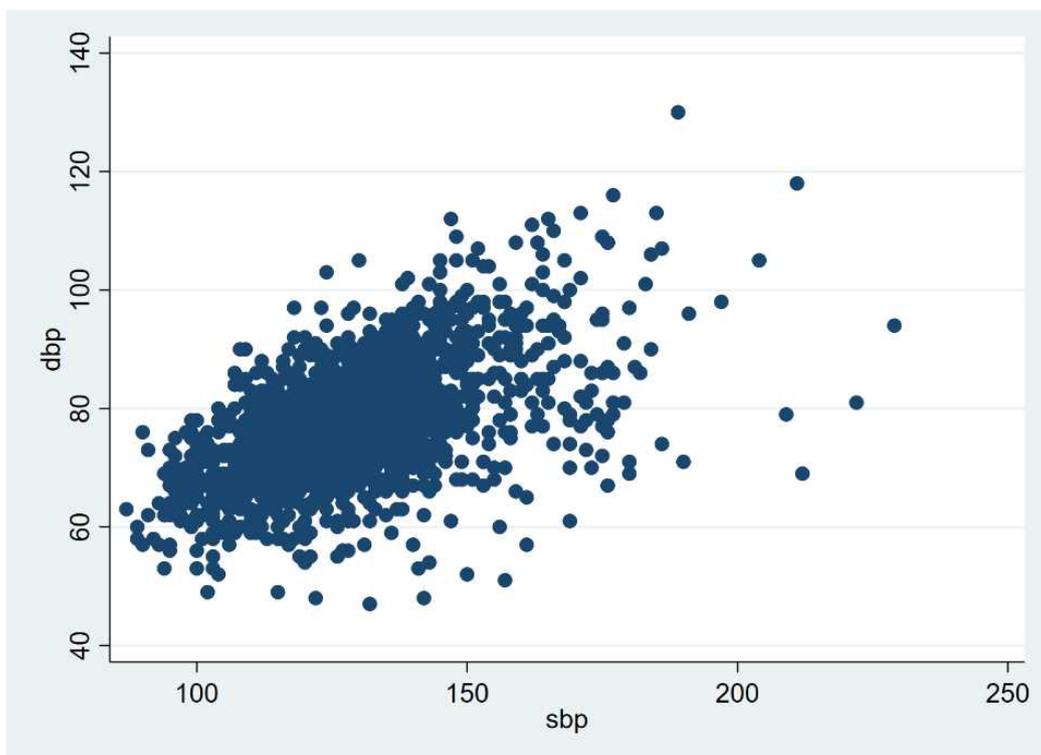
```
twoway
```

3.3 基本コマンド+オプション1

Stataの **twoway** コマンドでは、**twoway** の後に描くグラフを指定します。ここでは散布図を描画するサブコマンド **scatter** を指定しています。他のグラフはコマンドウィンドウに **help twoway** と入力することで調べることが出来ます。なお、サブコマンドが1つのみ場合は、**twoway** を省略可能です。

drGraph_Part1.do

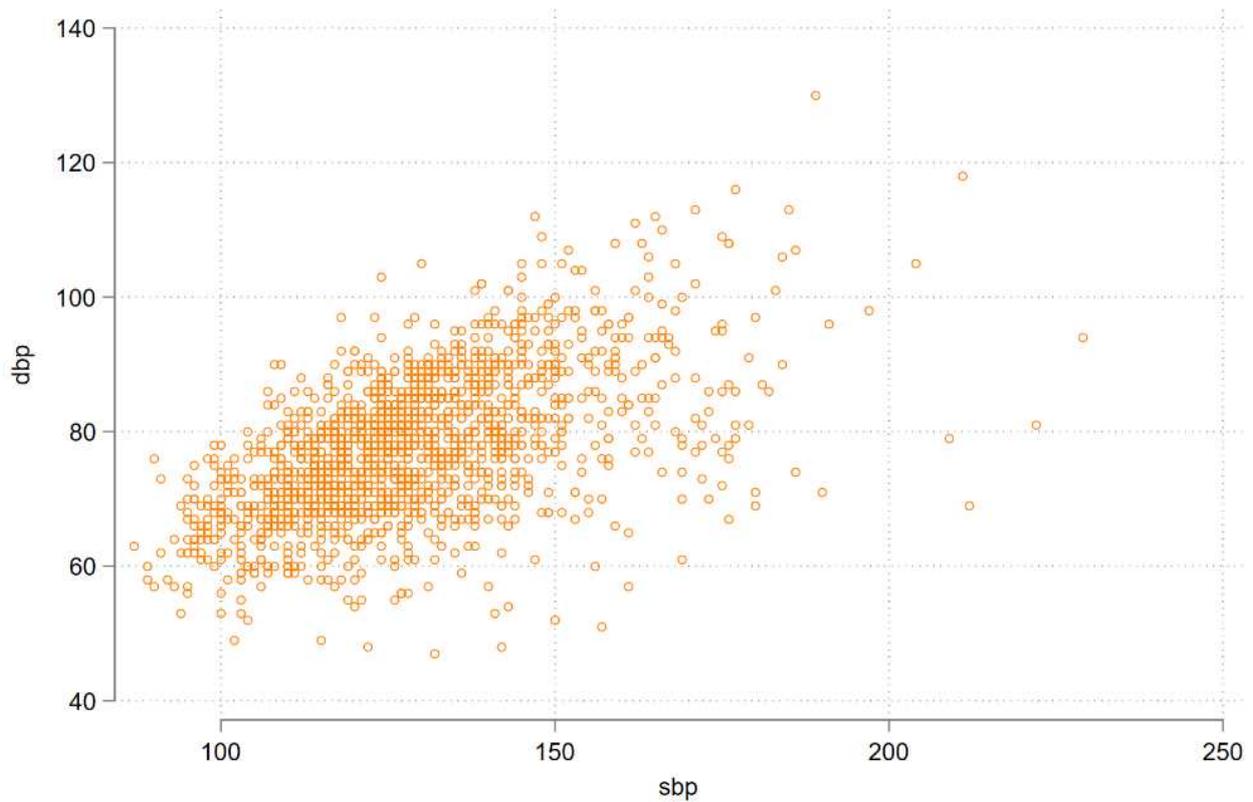
```
twoway scatter sbp dbp
```



3.4 基本コマンド+オプション 2

描いたプロットについて、Stata では、コマンドオプション（コンマの後に指定する）で編集します。ここでは3つのオプションを指定しています。

```
drGraph_Part1.do  
twoway scatter dbp sbp, ///  
    scheme(white_tableau) ///  
    msymbol(oh)          ///  
    mcolor(orange %30)   ///  
    // プロット色と透明度
```

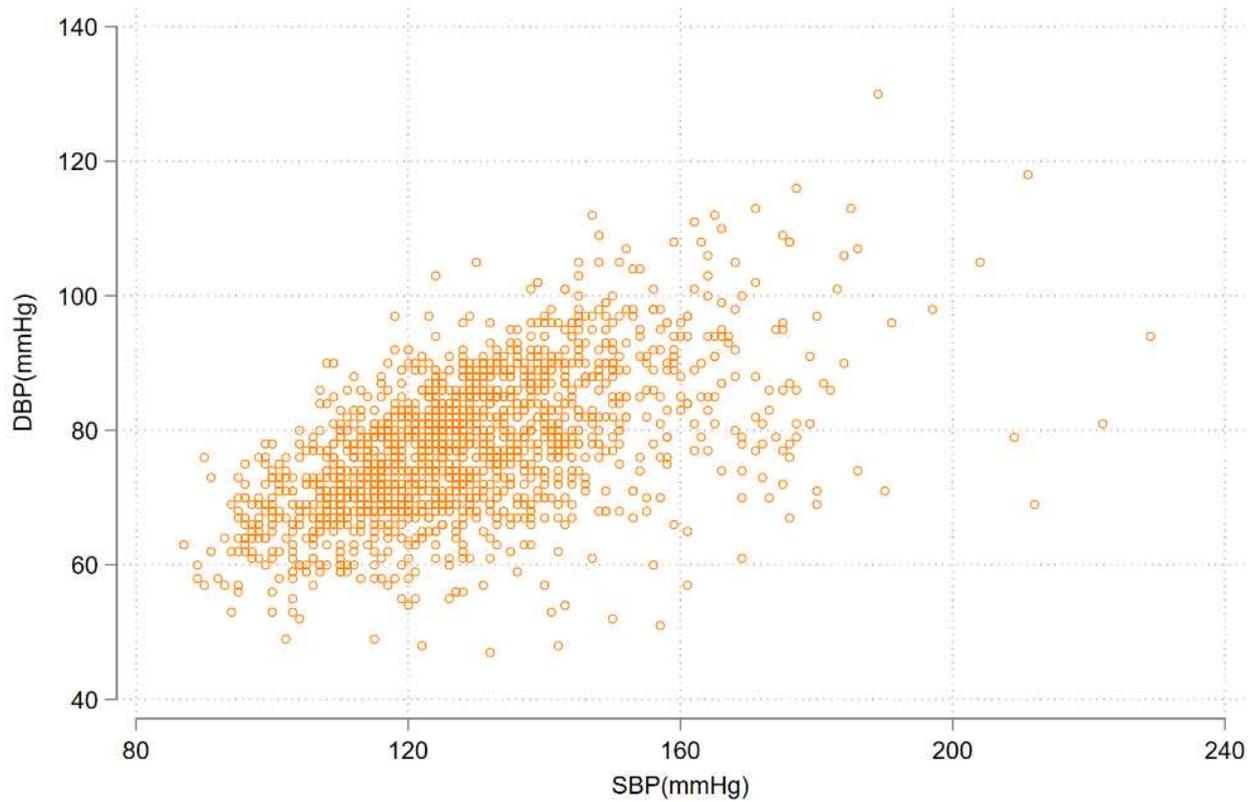


3.5 基本コマンド+オプション 3

Stata のコマンドのオプションをさらに付け加え、軸に関する情報を増やしました。

```
part1.do  
twoway scatter dbp sbp, ///  
    scheme(white_tableau) ///  
    msymbol(oh)          ///  
    mcolor(orange %30)   ///  
    // プロット色と透明度
```

```
xtitle("SBP(mmHg)")    /// X 軸
ytitle("DBP(mmHg)")    /// Y 軸
xlabel(80(40)240)       /// X 軸の範囲
ylabel(40(20)140)      /// Y 軸の範囲
```



3 tidy なデータ作成について

Stata では R の **tidyverse** に相当するようなパッケージ群は組み立てられていないようです。