Format\_Output

回归结束需要格式化输出，常见的表格输出主要有三种：

1. 描述性统计
2. 相关系数矩阵
3. 回归结果

解决方案统一用stargazer包，语法也比较统一。

这比Stata分别用logout,outreg2方便的多，结果也更美观。

下面的示例还是用样本选择模型自带的数据，重点演示如何输出。

stargazer支持输出到txt,latex,html。

html可以直接复制到excel或word里再稍微调整即可。

require(sampleSelection)

## Loading required package: sampleSelection  
## Loading required package: maxLik  
## Loading required package: miscTools  
##   
## Please cite the 'maxLik' package as:  
## Henningsen, Arne and Toomet, Ott (2011). maxLik: A package for maximum likelihood estimation in R. Computational Statistics 26(3), 443-458. DOI 10.1007/s00180-010-0217-1.  
##   
## If you have questions, suggestions, or comments regarding the 'maxLik' package, please use a forum or 'tracker' at maxLik's R-Forge site:  
## https://r-forge.r-project.org/projects/maxlik/

data(Mroz87)  
Mroz87$kids <- ( Mroz87$kids5 + Mroz87$kids618 > 0 )  
  
  
# Heckman两阶段估计  
heck<- heckit( lfp ~ age + I( age^2 ) + faminc + kids + educ,  
 wage ~ exper + I( exper^2 ) + educ + city+ kids, data=Mroz87 )  
  
# 极大似然估计  
ml<-selection( lfp ~ age + I( age^2 ) + faminc + kids + educ,  
 wage ~ exper + I( exper^2 ) + educ + city+ kids, data=Mroz87 )   
  
  
  
m <- selection(lfp ~ educ + age + kids5 + kids618 + nwifeinc,  
 wage >= 5 ~ educ, data = Mroz87 )

自定义了个corr函数来输出相关系数矩阵，下三角是Pearson上三角Spearman。

看了下SAS和Stata相关code感觉还是R简单明了。

corr(x)只接收一个数据框或者矩阵参数。

corr<-function(x){  
 p\_cor<-cor(x);  
 s\_cor<-cor(x,method="spearman");  
 p\_cor[upper.tri(p\_cor)==TRUE]<-s\_cor[upper.tri(s\_cor)==TRUE];  
 return(p\_cor);  
}

分别输出描述性统计，相关系数矩阵和回归结果。

stargazer对dplyr支持的很好，直接就能输出select的结果。

#描述性统计输出，summary.logical控制是否输出逻辑型变量  
  
stargazer(select(Mroz87,lfp,educ,age,faminc,kids,city),  
 type="html",title="Descriptive Statistics",  
 out.header=TRUE,summary.logical=FALSE,  
 digits=2,median=TRUE,  
 out="d:/descriptives.htm"  
 )

##   
## <table style="text-align:center"><caption><strong>Descriptive Statistics</strong></caption>  
## <tr><td colspan="7" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">Statistic</td><td>N</td><td>Mean</td><td>St. Dev.</td><td>Min</td><td>Median</td><td>Max</td></tr>  
## <tr><td colspan="7" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">lfp</td><td>753</td><td>0.57</td><td>0.50</td><td>0</td><td>1</td><td>1</td></tr>  
## <tr><td style="text-align:left">educ</td><td>753</td><td>12.29</td><td>2.28</td><td>5</td><td>12</td><td>17</td></tr>  
## <tr><td style="text-align:left">age</td><td>753</td><td>42.54</td><td>8.07</td><td>30</td><td>43</td><td>60</td></tr>  
## <tr><td style="text-align:left">faminc</td><td>753</td><td>23,081.00</td><td>12,190.00</td><td>1,500</td><td>20,880</td><td>96,000</td></tr>  
## <tr><td style="text-align:left">city</td><td>753</td><td>0.64</td><td>0.48</td><td>0</td><td>1</td><td>1</td></tr>  
## <tr><td colspan="7" style="border-bottom: 1px solid black"></td></tr></table>

#相关系数矩阵输出  
stargazer(corr(select(Mroz87,lfp,educ,age,faminc)),  
 type="html",title="Correlations",  
 out.header=TRUE,summary.logical=FALSE,  
 digits=2,median=TRUE,  
 out="d:/cor.htm",  
 notes="Pearson correlation lower block and Spearman correlation upper block"  
)

##   
## <table style="text-align:center"><caption><strong>Correlations</strong></caption>  
## <tr><td colspan="5" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"></td><td>lfp</td><td>educ</td><td>age</td><td>faminc</td></tr>  
## <tr><td colspan="5" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">lfp</td><td>1</td><td>0.18</td><td>-0.08</td><td>0.15</td></tr>  
## <tr><td style="text-align:left">educ</td><td>0.19</td><td>1</td><td>-0.10</td><td>0.40</td></tr>  
## <tr><td style="text-align:left">age</td><td>-0.08</td><td>-0.12</td><td>1</td><td>0.07</td></tr>  
## <tr><td style="text-align:left">faminc</td><td>0.10</td><td>0.36</td><td>0.05</td><td>1</td></tr>  
## <tr><td colspan="5" style="border-bottom: 1px solid black"></td></tr><tr><td colspan="5" style="text-align:left">Pearson correlation lower block and Spearman correlation upper block</td></tr>  
## </table>

#ml,m,heck为模型变量，html输出方便复制到excel里调整  
#vc\*p代表星号放在coef上，如果vcp\*星号就放P值上了，无report选项默认汇报标准误  
#omit用来控制不显示特定的自变量，通常是dummy，omit.labels的长度得和omit数一致  
stargazer(ml,m,heck,type="html",title="selection regression table",  
 report="vc\*p",digits=3,  
 dep.var.labels=c("wage","wage0-1","wages"),  
 model.names=FALSE,model.numbers=TRUE,  
 column.labels=c("ML","Heck"),column.separate=c(2,1),  
 header=FALSE,  
 omit="kids",omit.labels="kids dummy",  
 flip=TRUE,out.header=TRUE,out="d:/reg.htm")

##   
## <table style="text-align:center"><caption><strong>selection regression table</strong></caption>  
## <tr><td colspan="4" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"></td><td colspan="3"><em>Dependent variable:</em></td></tr>  
## <tr><td></td><td colspan="3" style="border-bottom: 1px solid black"></td></tr>  
## <tr><td style="text-align:left"></td><td>wage</td><td>wage0-1</td><td>wages</td></tr>  
## <tr><td style="text-align:left"></td><td colspan="2">ML</td><td>Heck</td></tr>  
## <tr><td style="text-align:left"></td><td>(1)</td><td>(2)</td><td>(3)</td></tr>  
## <tr><td colspan="4" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">exper</td><td>0.027</td><td></td><td>0.023</td></tr>  
## <tr><td style="text-align:left"></td><td>p = 0.659</td><td></td><td>p = 0.709</td></tr>  
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>  
## <tr><td style="text-align:left">I(exper2)</td><td>-0.0003</td><td></td><td>-0.0001</td></tr>  
## <tr><td style="text-align:left"></td><td>p = 0.878</td><td></td><td>p = 0.955</td></tr>  
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>  
## <tr><td style="text-align:left">educ</td><td>0.464<sup>\*\*\*</sup></td><td>0.288<sup>\*\*\*</sup></td><td>0.439<sup>\*\*\*</sup></td></tr>  
## <tr><td style="text-align:left"></td><td>p = 0.000</td><td>p = 0.000</td><td>p = 0.00005</td></tr>  
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>  
## <tr><td style="text-align:left">city</td><td>0.436</td><td></td><td>0.436</td></tr>  
## <tr><td style="text-align:left"></td><td>p = 0.168</td><td></td><td>p = 0.168</td></tr>  
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>  
## <tr><td style="text-align:left">Constant</td><td>-1.861</td><td>-4.521<sup>\*\*\*</sup></td><td>-1.303</td></tr>  
## <tr><td style="text-align:left"></td><td>p = 0.142</td><td>p = 0.000</td><td>p = 0.541</td></tr>  
## <tr><td style="text-align:left"></td><td></td><td></td><td></td></tr>  
## <tr><td colspan="4" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">kids dummy</td><td>Yes</td><td>No</td><td>Yes</td></tr>  
## <tr><td colspan="4" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left">Observations</td><td>753</td><td>753</td><td>753</td></tr>  
## <tr><td style="text-align:left">R<sup>2</sup></td><td></td><td></td><td>0.127</td></tr>  
## <tr><td style="text-align:left">Adjusted R<sup>2</sup></td><td></td><td></td><td>0.115</td></tr>  
## <tr><td style="text-align:left">Log Likelihood</td><td>-1,581.000</td><td>-653.200</td><td></td></tr>  
## <tr><td style="text-align:left">rho</td><td>-0.104 (0.192)</td><td>0.116 (0.271)</td><td>-0.236</td></tr>  
## <tr><td style="text-align:left">Inverse Mills Ratio</td><td></td><td></td><td>-0.741 (1.411)</td></tr>  
## <tr><td colspan="4" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"><em>Note:</em></td><td colspan="3" style="text-align:right"><sup>\*</sup>p<0.1; <sup>\*\*</sup>p<0.05; <sup>\*\*\*</sup>p<0.01</td></tr>  
## </table>