

CONCEPTUAL INTRODUCTION TO SAS

1 minute Why SAS

1 minute SAS Windows

8 minutes 10 SAS Concepts

5 minutes SAS live

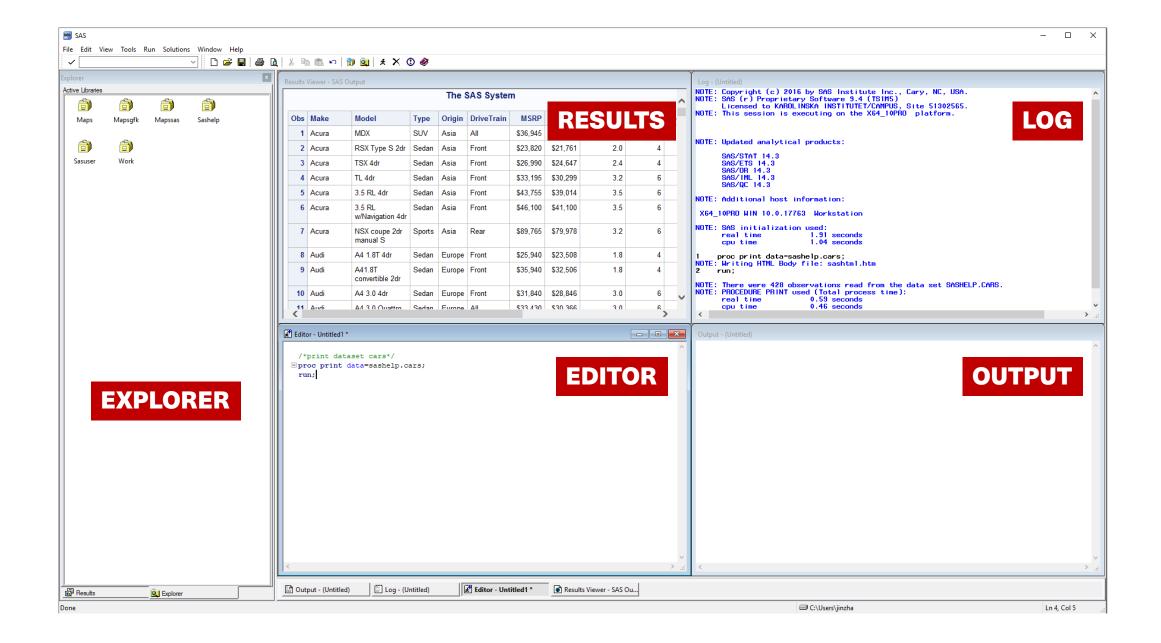
Top 3 SAS features

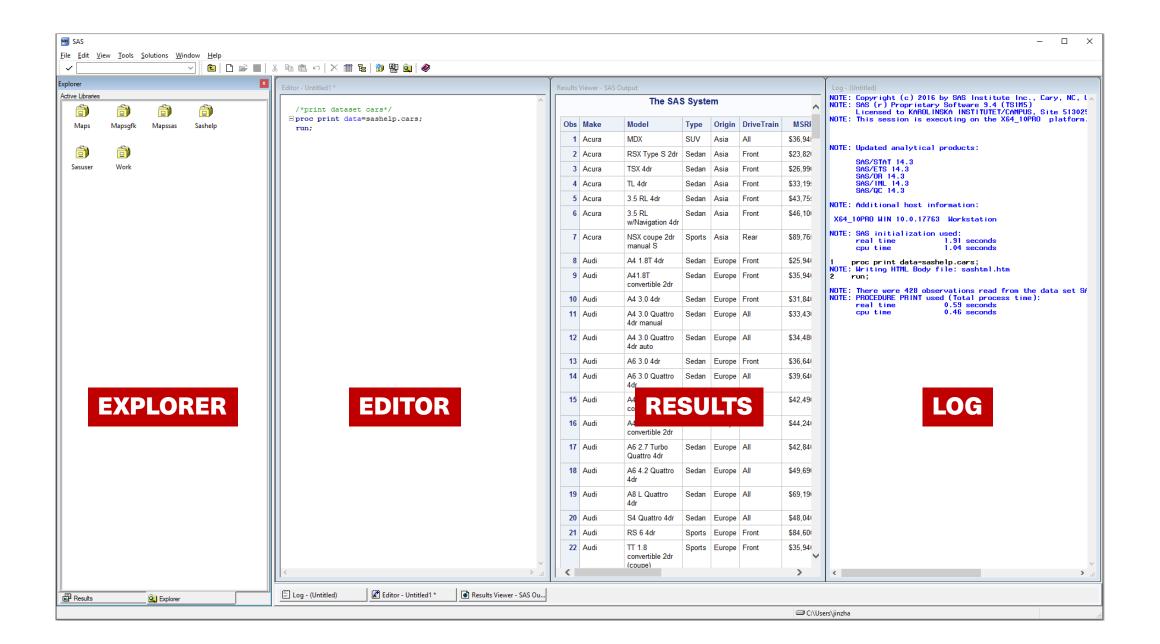
1. SQL + SAS = All-in-one data management

2. Syntax reads (almost) like English

3. Robust and backwards compatible

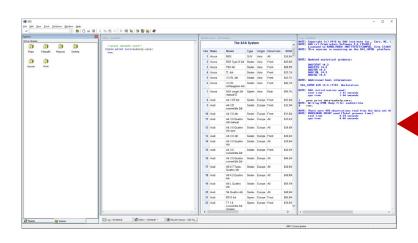
SAS WINDOWS





#1 SAS IDEs

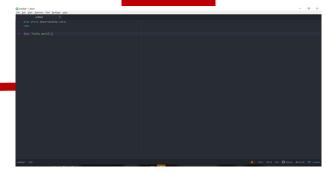
BASE SAS



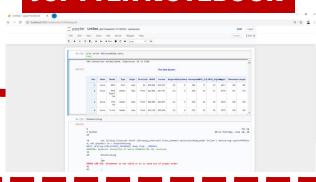
SAS ENTERPRISE GUIDE



ATOM



JUPYTER NOTEBOOK

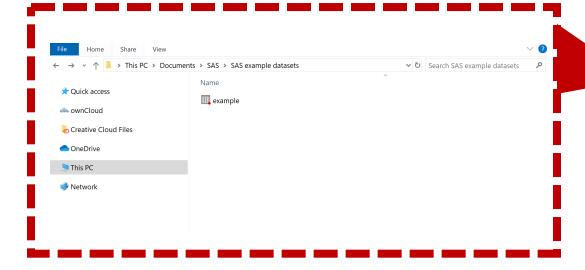


#2 SAS-logic is "row-by-row"

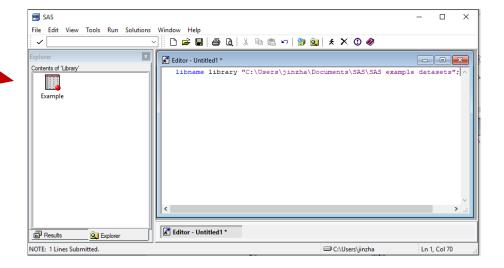


#3 You don't load datafiles into memory

PHYSICAL FOLDER

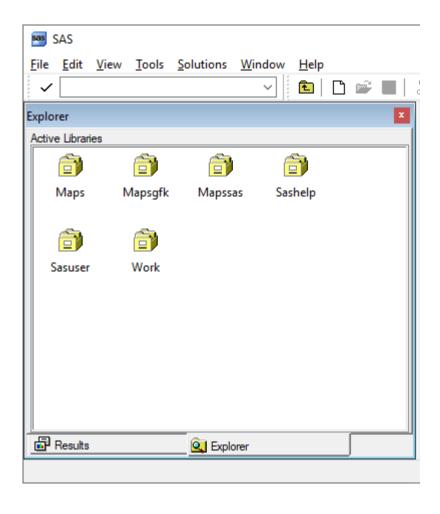


BASE SAS



#3 Work Folder

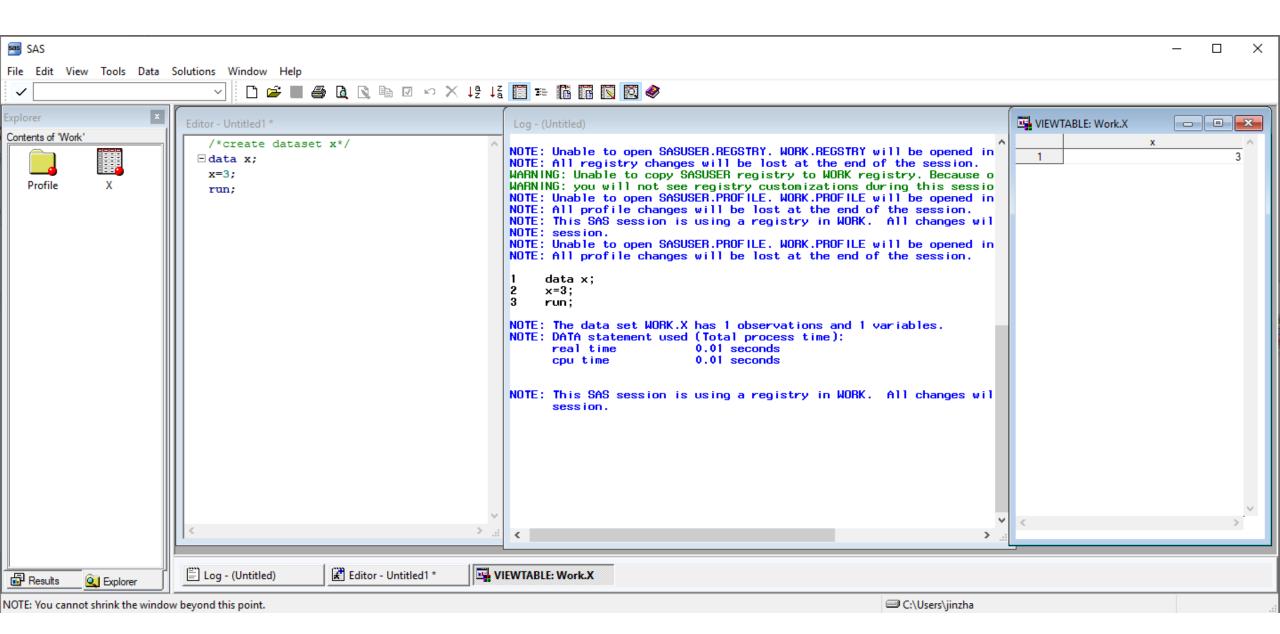
#3 Work Folder

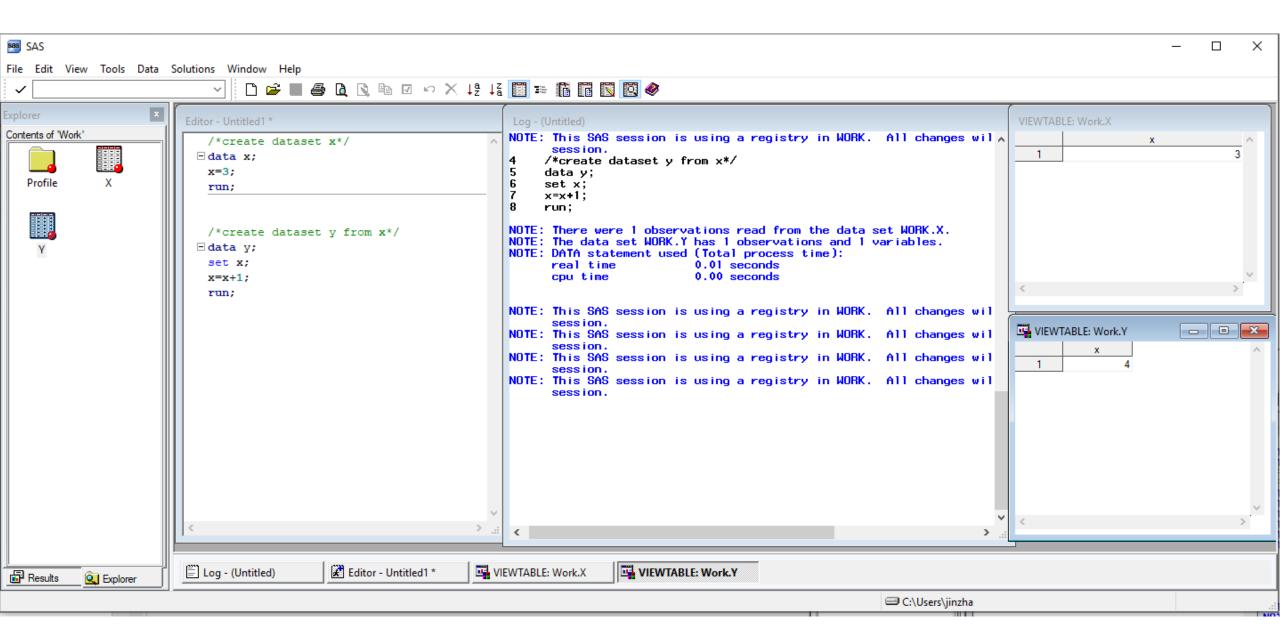


Temporary folder
 (deletes after each session)

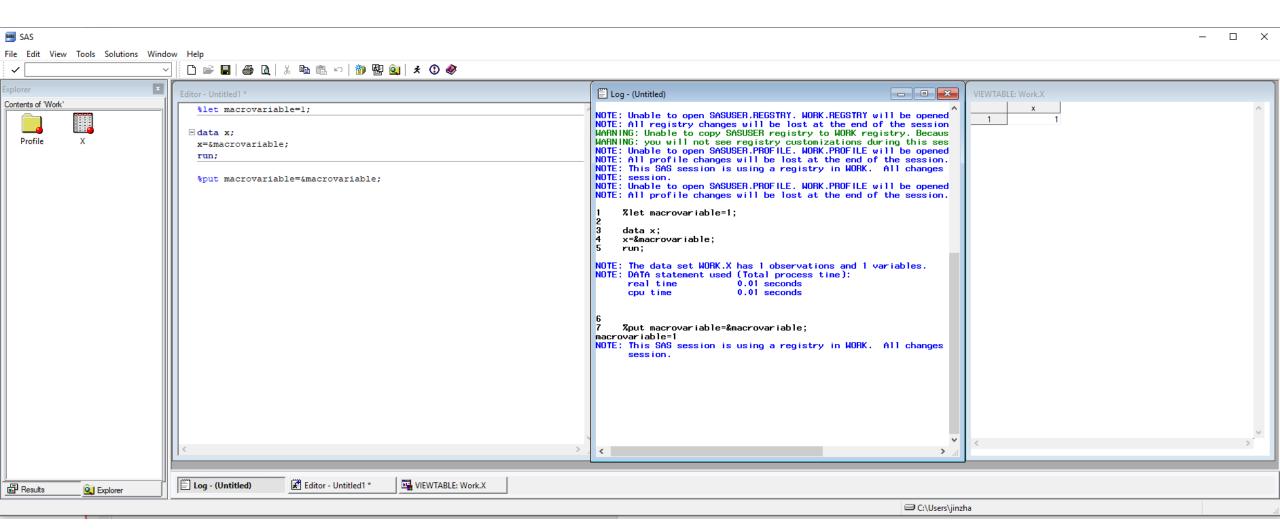
 By default datasets are created in work folder

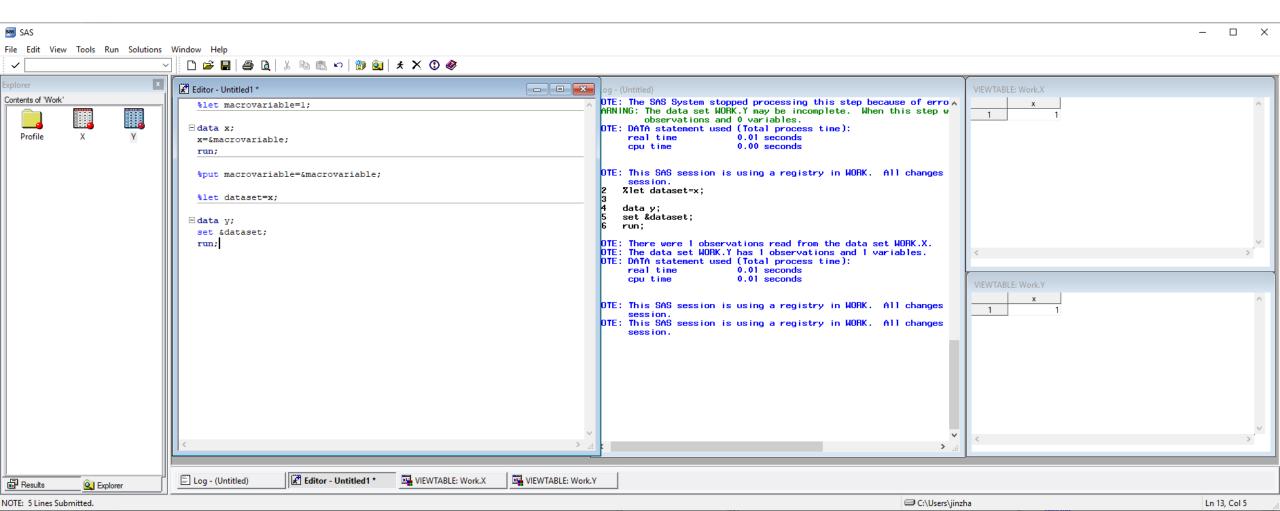
#4 Datasets





#5 Macro Variables





#6 Variable Types

Only 2 variable types

Character

Numerical

String variables

Floats and Integers

Thousands of built-in formats

Character

Varying lengths \$customformats.

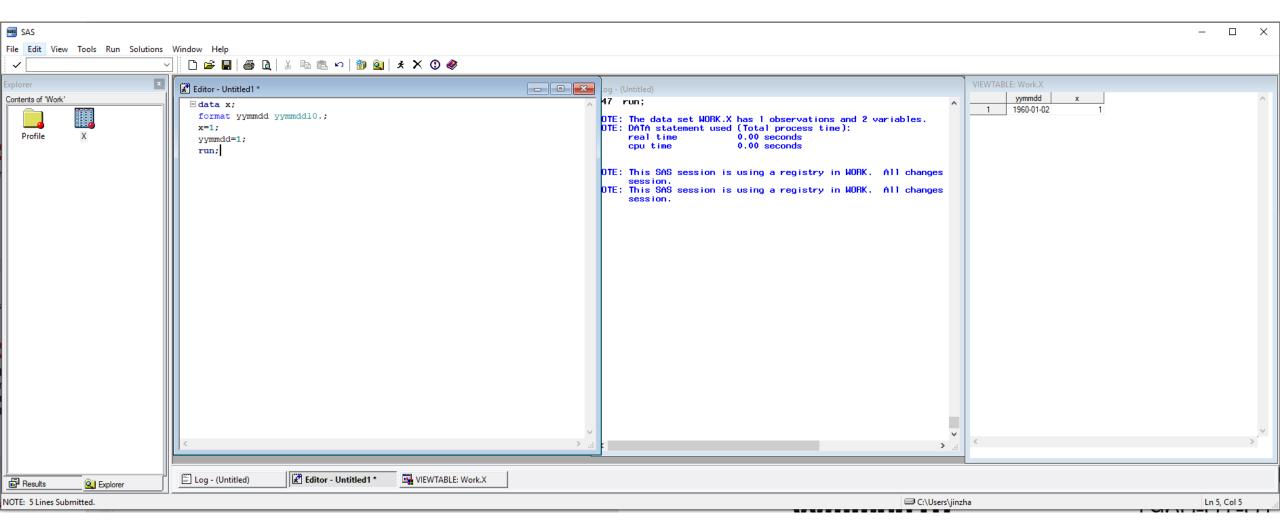
Numerical

datetime20. yymmdd10. hhmm5. 8.2

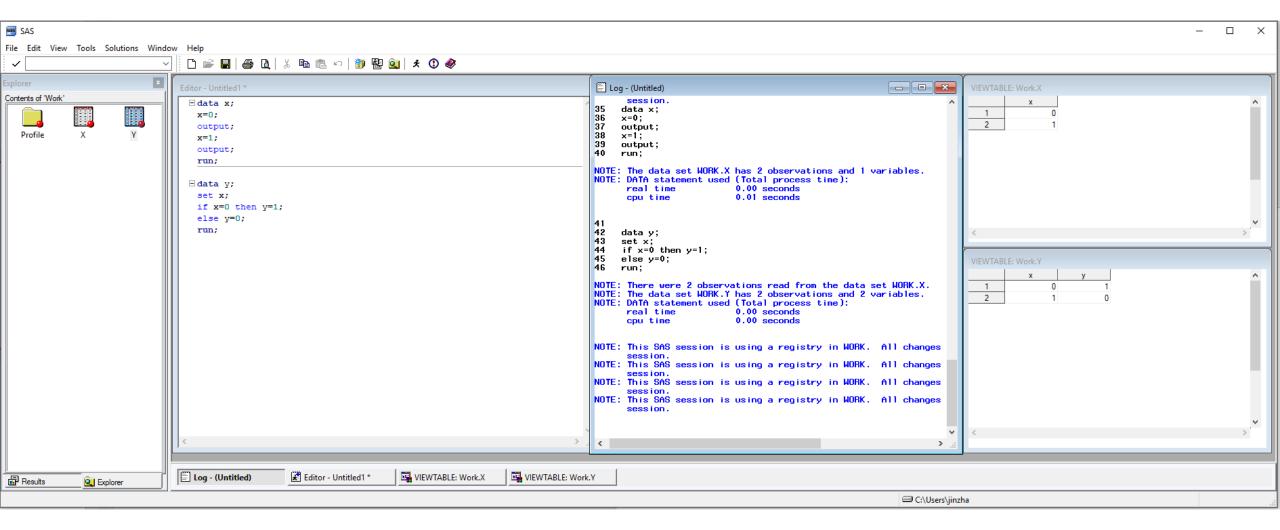
Thousands of built-in formats

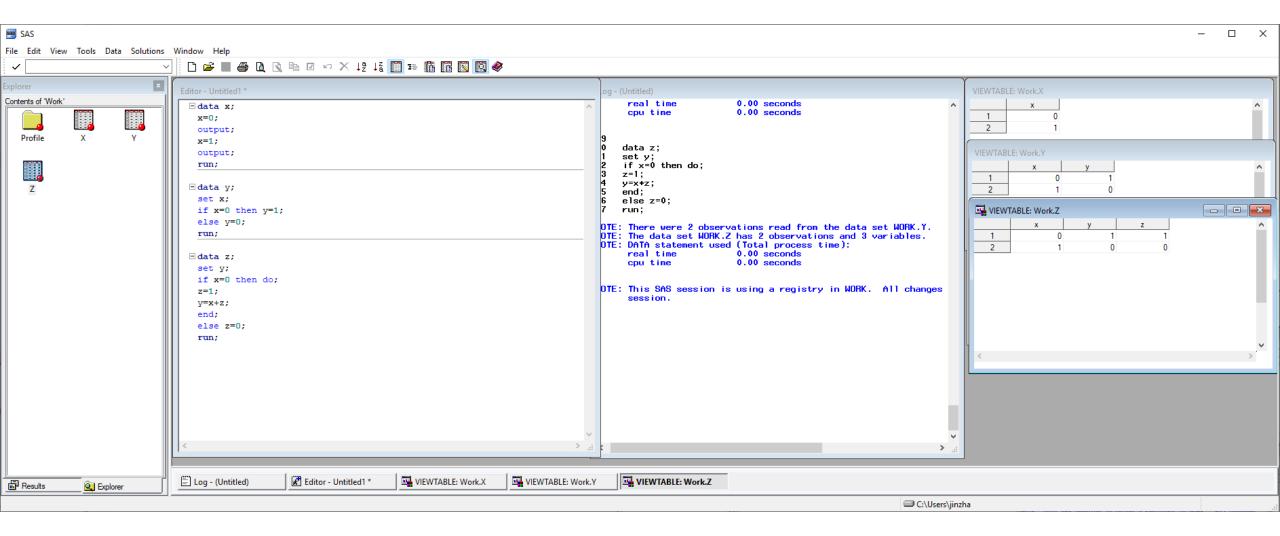
Numerical

datetime20. yymmdd10. hhmm5. 8.2 02JAN1960:06:00:00 1960-01-02 06:00 1.00

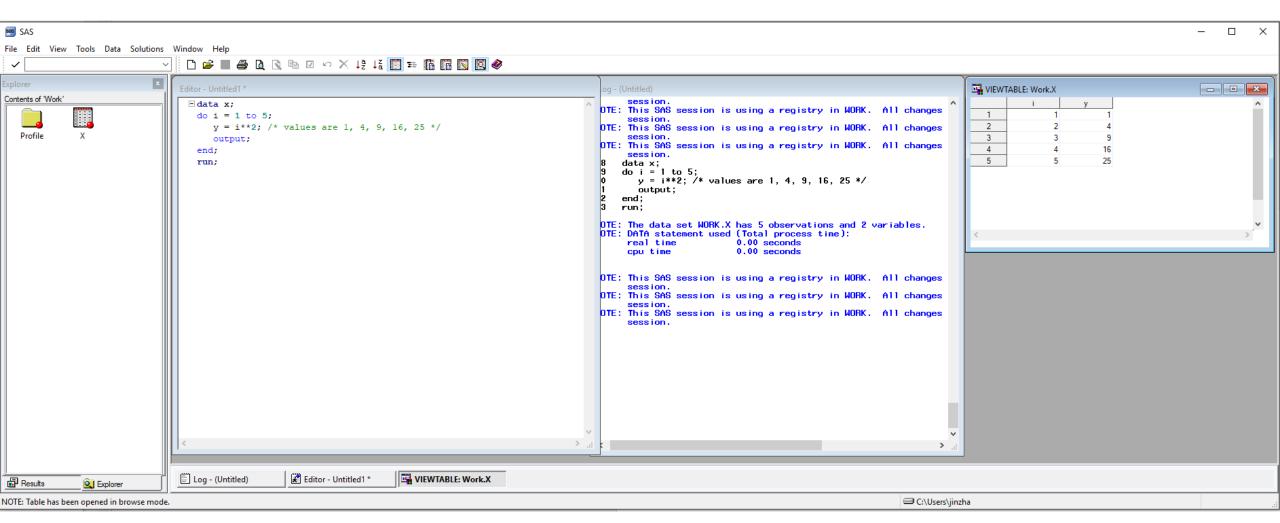


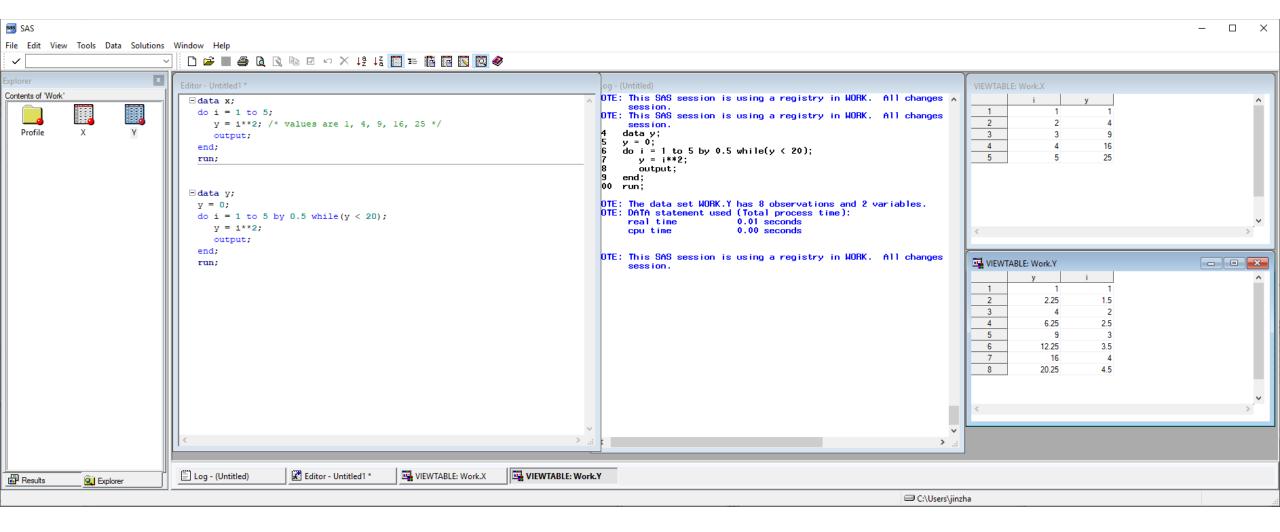
#7 If-then clauses

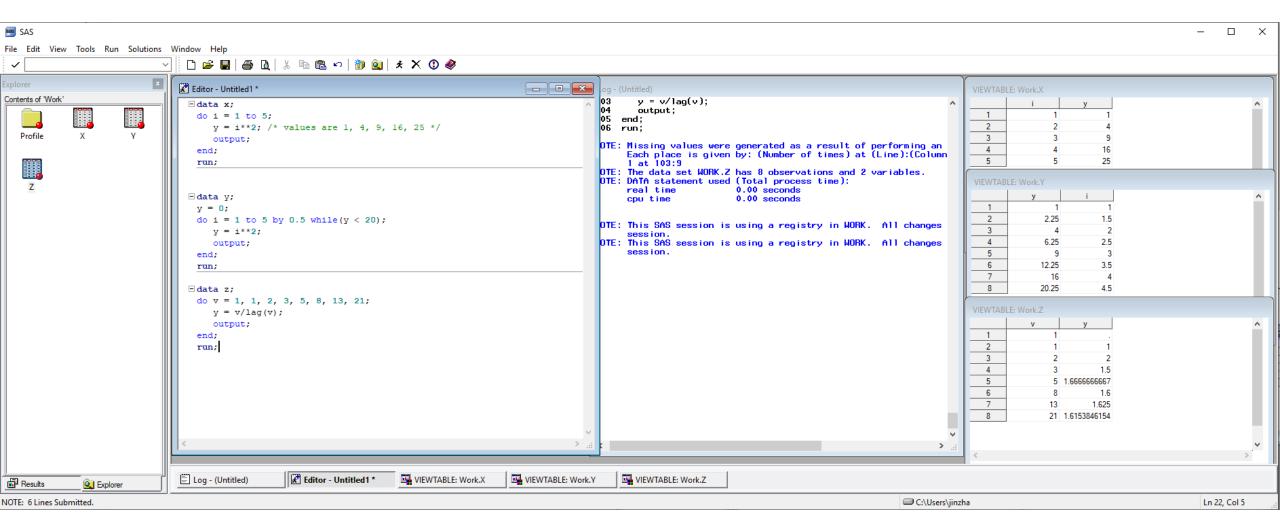




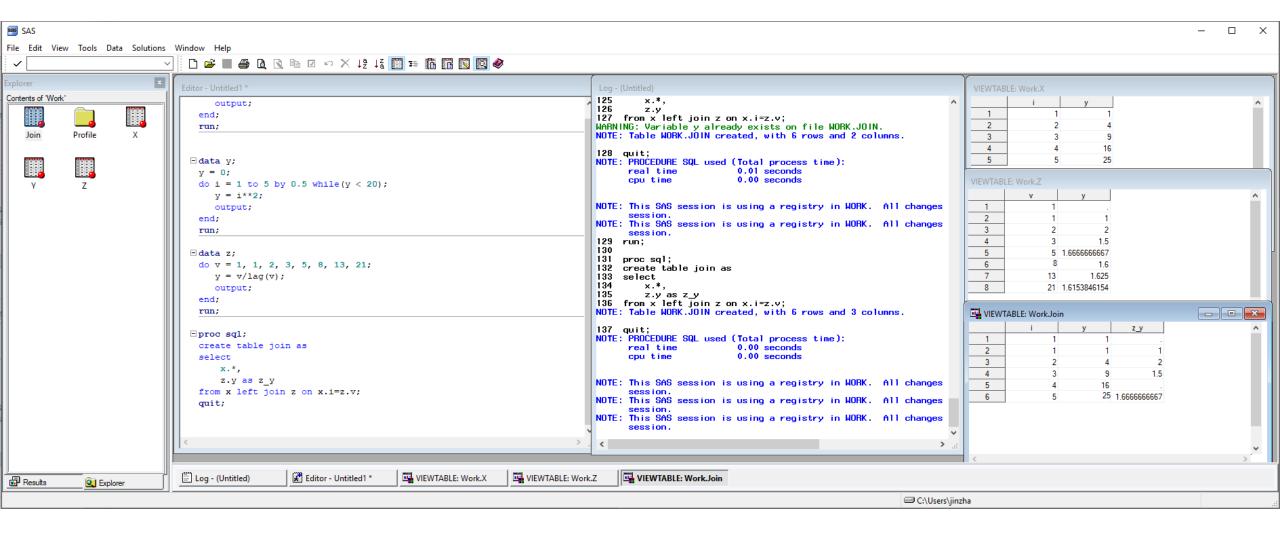
#8 Loops



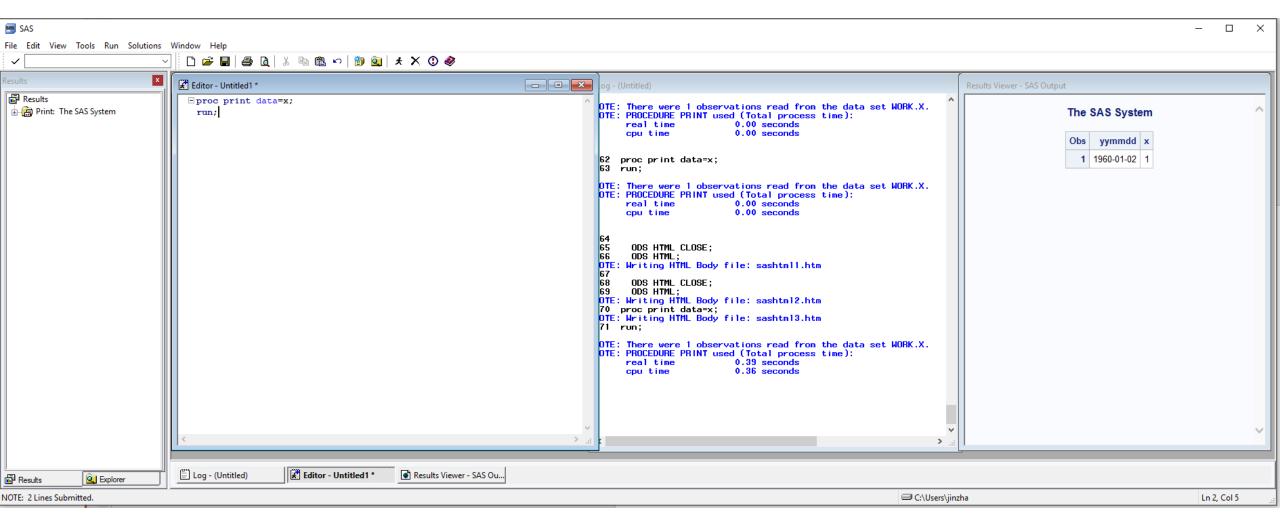


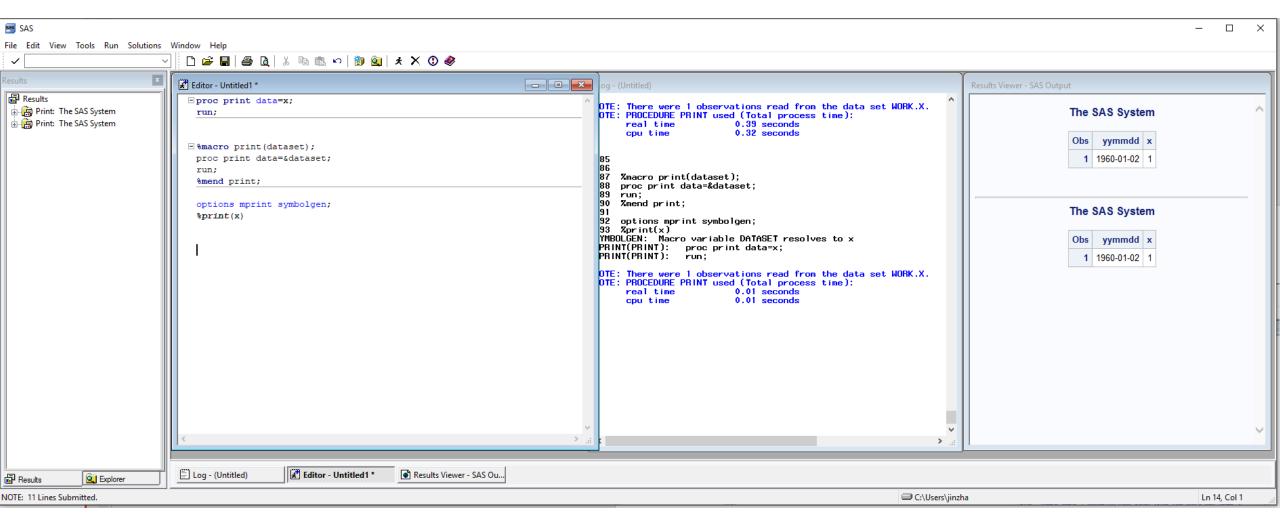


#9 Proc SQL

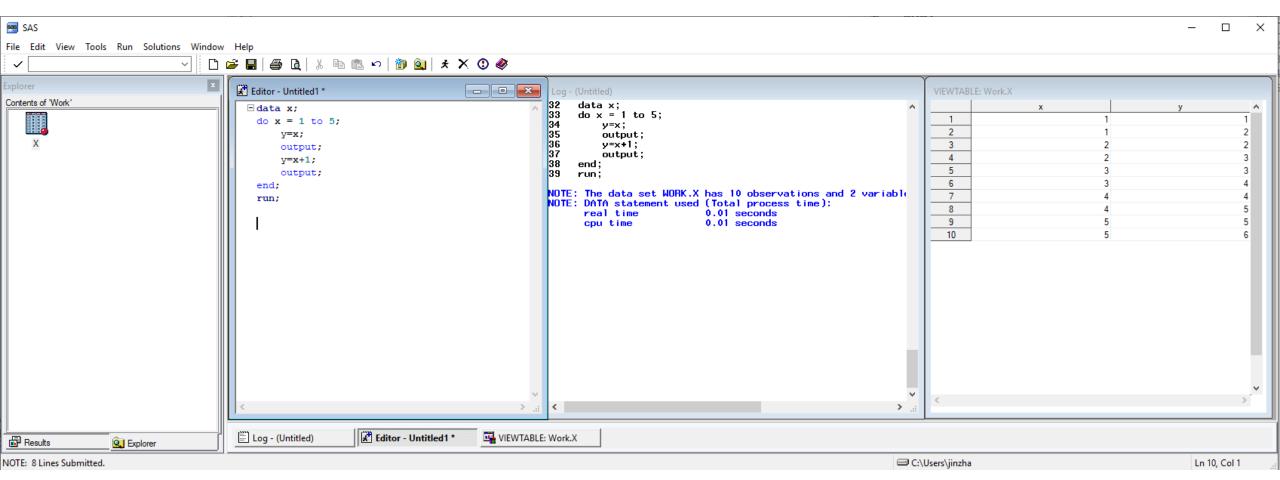


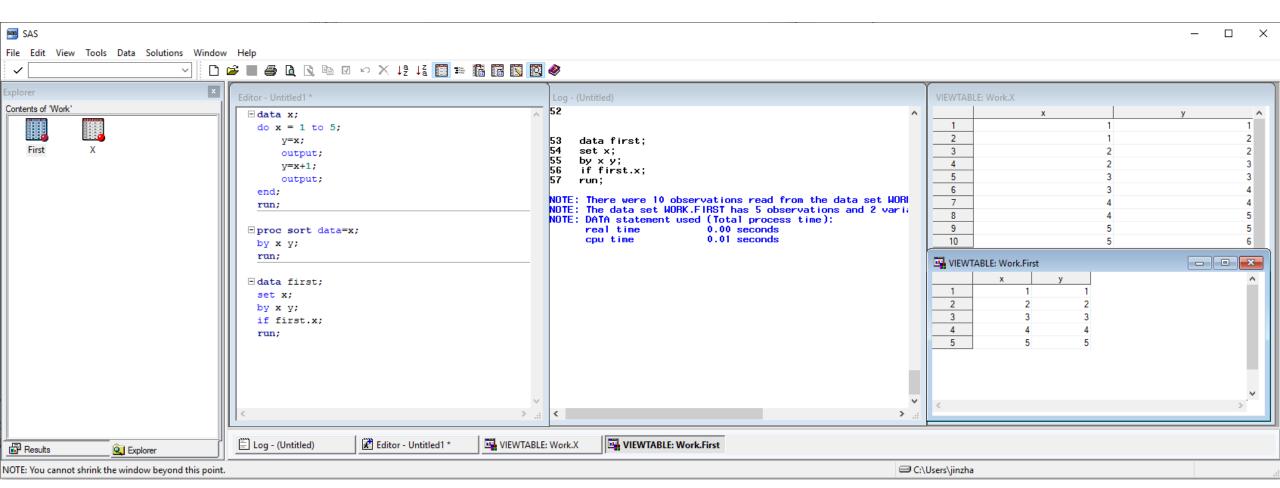
#10 Macros





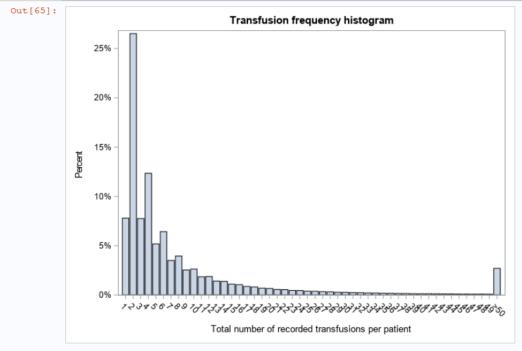
Bonus #1 Processing "BY-groups"





Bonus #2 Custom formats

```
In [65]: proc sql;
         create table p_tx as
         select distinct
             a.idnr,
              count (a.idnr) as transfusions
         from transfusion a inner join clean.persons b on a.idnr=b.idnr where b.error=0
         group by a.idnr;
         create table p_tx_freq as
         select distinct
             transfusions,
             count (transfusions) as count
         from p tx
         group by transfusions;
         quit;
         proc format;
         value fiftyplus
          50-high = ">50"
         title "Transfusion frequency histogram";
         proc sgplot data=p_tx_freq;
         format transfusions fiftyplus.;
         vbar transfusions / response=count stat=percent;
         yaxis label="Percent";
         xaxis label="Total number of recorded transfusions per patient";
         run;
```



SAS CONCEPTS RECAP

```
#1 Choose your favorite IDE
#2 Row-by-row logic (magnet bands)
#3 Work folder – default and temporary folder
#4 Create datasets with the "data" step
#5 Create macro variables with "%let"
#6 Char + Num variables, but many formats!
#7 If-then-else
#8 Loops – do, until, while
#9 Native support for SQL in Proc SQL
#10 Use SAS Macros to save time
```

Bonus #1 Use "by processing" and first.variables Bonus #2 Use custom formats

Demo