Financial Data on WRDS :
Contents, Analysis and Methods of Access

Mark Keintz
Senior Data Analyst

wrds WHARTON RESEARCH DATA SERVICES
The Global Standard for Business Research

Agenda

1. What is WRDS?

2. Accessing the Data

3. Sources of Financial Data

4. The Big Two: Compustat and CRSP

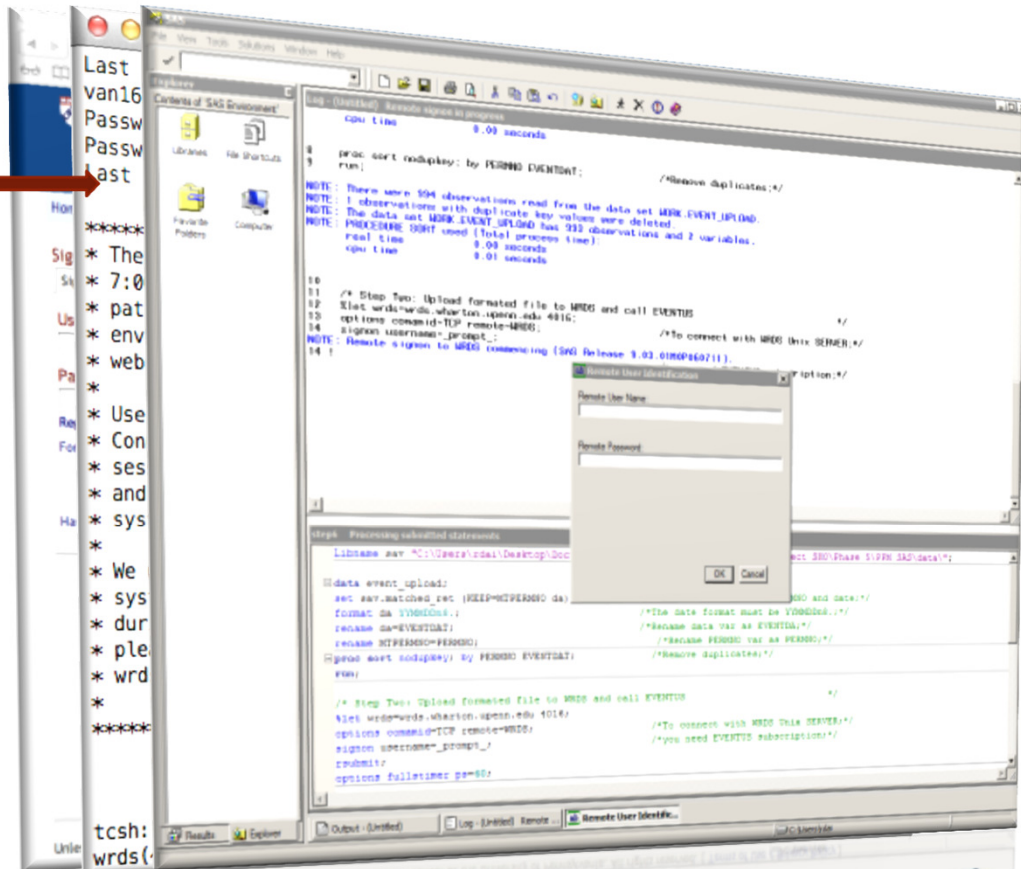5. Advanced Access Features/ Using SAS at WRDS

wrds

WRDS Overview

+ Single source for leading business research databases
  + Accounting, Stocks, Bonds, Loans, Banking, Insurance, Real Estate, Economics, Marketing, Statistics, and +
  + Standardized queries: Query + Variable Documentation + List of Manuals, & Research Overviews and Applications
  + Linkage across different data sources, e.g. accounting data with stock market data

+ Full-time research and technical support
  + Online help, email support, and 24/7 network monitoring
  + Data Overviews, Research Applications, and Knowledge Base

wrds
The Global Standard for Business Research
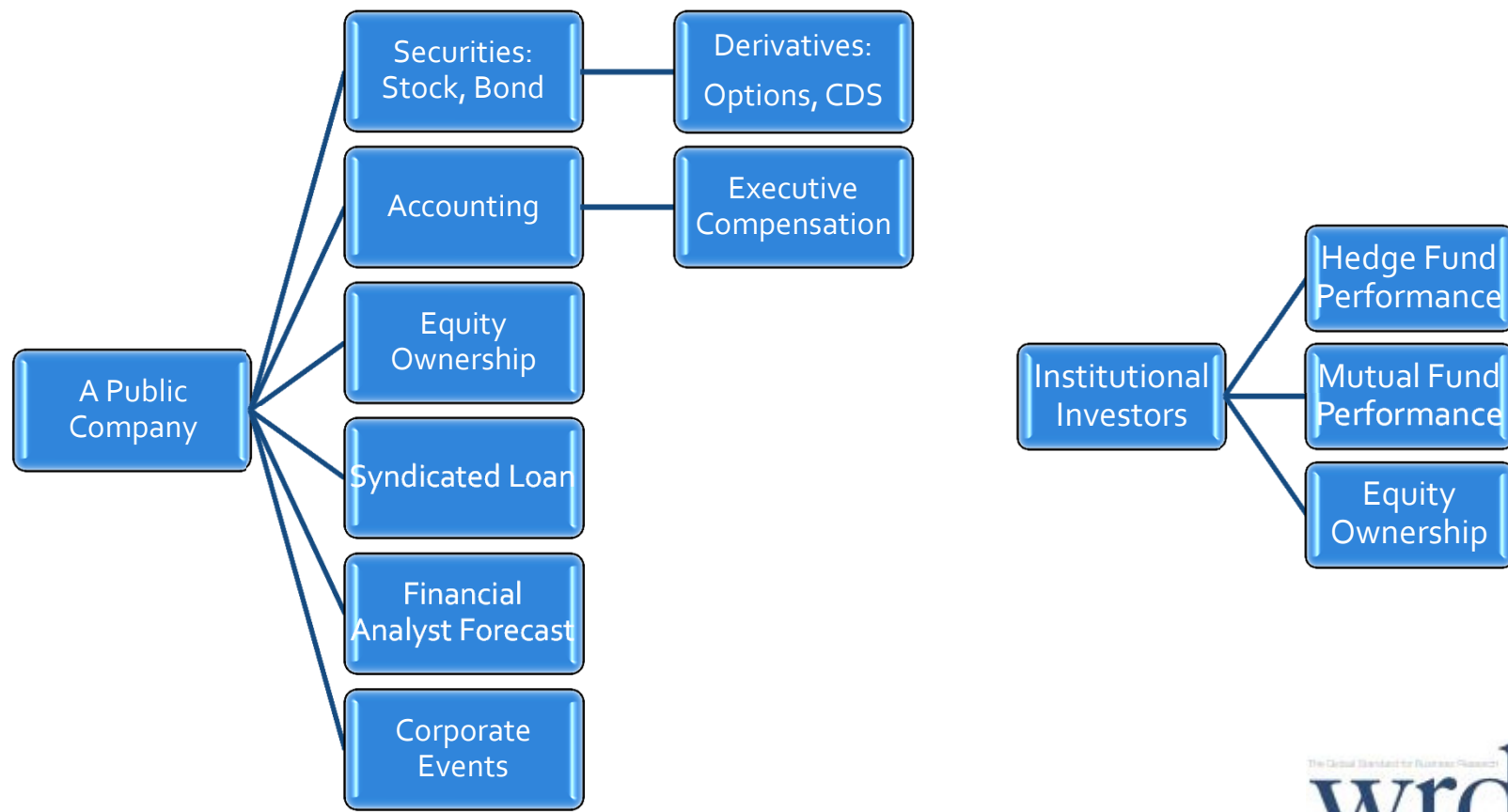
Multiple Access Options:

WRDS   Internet

- **Web Queries:** Compatible with Various Browsers and Operating Systems
- **Unix Logon:** Large Data Manipulation
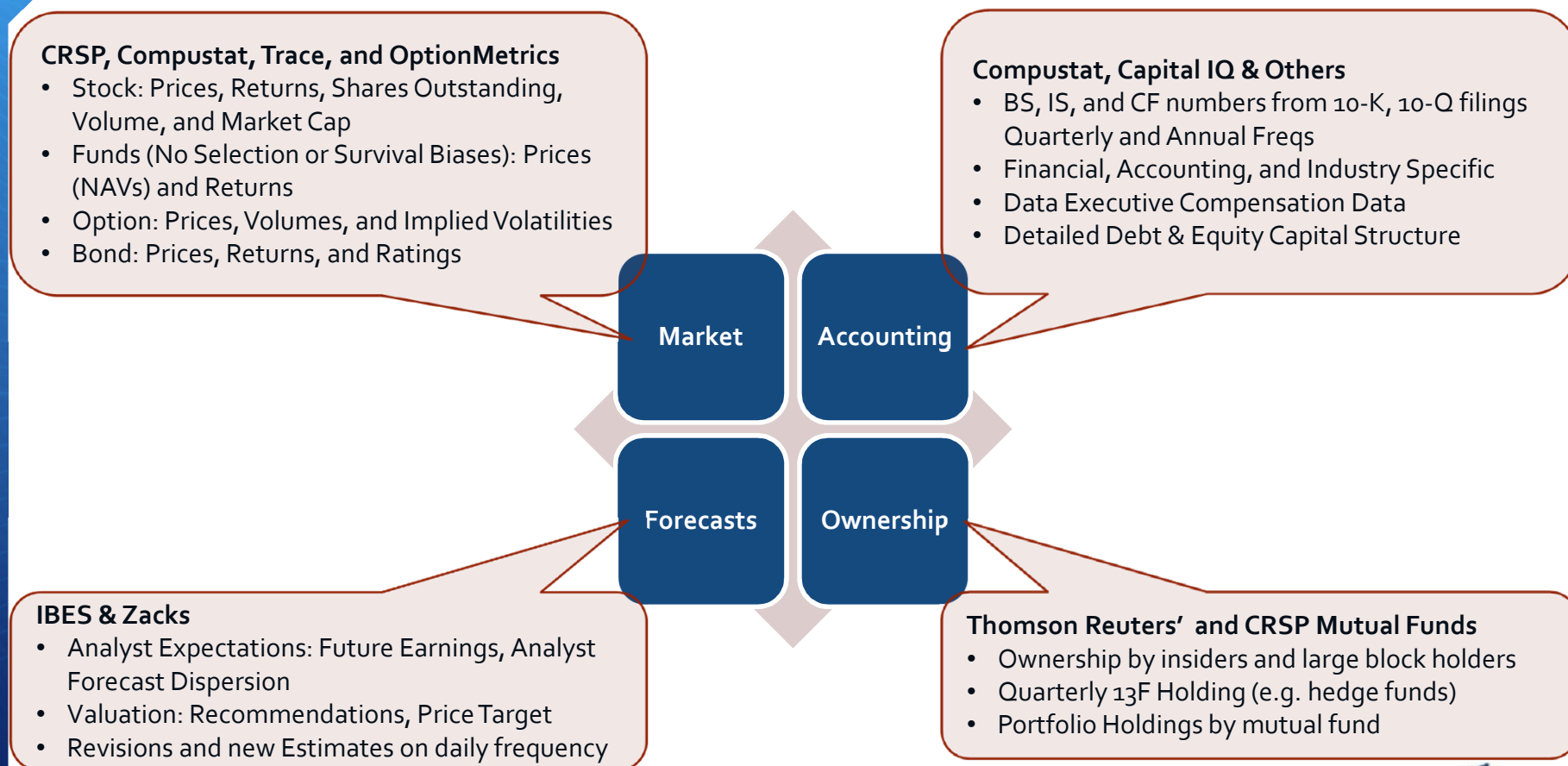- **PC SAS/CONNECT:** Interactive Interface

Financial Information by Source

Financial Information Integration

CRSP, Compustat, Trace, and OptionMetrics
- Stock: Prices, Returns, Shares Outstanding, Volume, and Market Cap
- Funds (No Selection or Survival Biases): Prices (NAVs) and Returns
- Option: Prices, Volumes, and Implied Volatilities
- Bond: Prices, Returns, and Ratings

Compustat, Capital IQ & Others
- BS, IS, and CF numbers from 10-K, 10-Q filings Quarterly and Annual Freqs
- Financial, Accounting, and Industry Specific
- Data Executive Compensation Data
- Detailed Debt & Equity Capital Structure

Market

Accounting

Forecasts

Ownership

IBES & Zacks
- Analyst Expectations: Future Earnings, Analyst Forecast Dispersion
- Valuation: Recommendations, Price Target
- Revisions and new Estimates on daily frequency

Thomson Reuters' and CRSP Mutual Funds
- Ownership by insiders and large block holders
- Quarterly 13F Holding (e.g. hedge funds)
- Portfolio Holdings by mutual fund

Compustat :Accounting and Financial Data

+ Compustat North-America: U.S. and Canadian fundamental and market information on 31,000+ active and inactive public companies, from 1950-present

+ Fundamental data available on an annual and quarterly frequencies with thousands of Income Statement, Balance Sheet, Statement of Cash Flows, and supplemental & industry-specific data items

+ Market data available on a monthly and daily frequencies with Prices, Dividends, Returns, Trading Volume , Shares Outstanding and Short-Interest Information (Very Similar to CRSP in many dimensions)

Compustat NA Example

+ Get Current Assets, Inventories and Long-Term Debt for Microsoft, Dell, and IBM between Jan 2000 and April 2009

WRDS Support: Compustat

+ Overview to Compustat XPressfeed Database

Support → Data → Manuals And Overviews → Compustat → General → Overview of Compustat Expressfeed

+ Sample Programs (extracts, point-in-time, etc)

Support → Data → Sample Programs → Compustat → Compustat Sample Programs

+ Research Applications (book-to-market, P/E Ratio)

Research → Applications → Wharton Research Data Services

wrds

CRSP: Stock Market Data

+ Center for Research in Security Prices (CRSP) is a research center at the Booth School of Business of the University of Chicago

+ Comprehensive collection of daily and monthly security price, return, and volume data for the NYSE/AMEX/NASDAQ/ARCA trading venues

+ Daily and Monthly data for roughly 28,000+ securities of Domestic companies and ADRs traded on major exchanges (NYSE/AMEX/Nasdaq/ARCA), from 1925–present

+ Complete historical information (bias-free):
  + Accurate records of special distributions and stock splits in return calculation
  + Keep delisting companies (pre-M&A or bankruptcies), and delisting returns

CRSP Example

+ **Get monthly prices, return and volume information for Microsoft and Ford from 1925 to 2008**

WRDS Support: CRSP

+ Overview of CRSP/Compustat Merged Database in WRDS

Support → Data → Manuals And Overviews → CRSP → CRSP Manuals and Overviews

+ Sample Programs

  + Data Extracts, CCM and merging by CUSIP

  + Calculate CAPM beta and excess returns

Support → Data → Sample Programs → CRSP → CRSP Sample Programs

+ Research Applications

  + Compound returns

  + Momentum and Governance Portfolios,

  + Fama-French factors and B/M portfolios,

  + Beta Estimation

Research → Applications → Wharton Research Data Services

Link CRSP and Compustat

+ **CRSP**: the data source for US security (equity) trading records.
  + since 1925, and 28,673 Securities (until 2010)

+ **Compustat**: the data source of accounting items for US public firms
  + since 1950 and 29,701 Companies (until 2010)

+ Major Difficulties to link CRSP and Compustat
  + Different Frequencies:
    + Compustat: Accounting items are on Annual or Quarter base
    + CRSP: Security trading data is Monthly or Daily
  + Different Universe of Companies
    + CRSP: Equities listed in NYSE, AMEX, NASDAQ and ARCA
    + Compustat: 10K and 10Q Filers to the SEC

Common Secondary Identifiers

+ **Stock Tickers**: Exchange Specific, with Suffixes, *NOT* Permanent, and Recyclable.

+ **CUSIP** also is *NOT* Permanent but Unique and *NOT* Exchange Specific.
    + 8- or 9-digit CUSIP is assigned for issues (e.g. equities and bonds).
    + 6-digit CUSIP is assigned for issuers (e.g. firms)

+ **CIK**: US filing number at the SEC, and Changeable

+ **Company Names** are often abbreviated/manipulated by data vendors
    + Very difficult to have a reliable merging code by using company names

wrds

Identifiers in CRSP

CRSP PERMANENT NUMBER	CRSP PERMANENT COMPANY NUMBER	START DATE OF EFFECTIVE NAME	END DATE OF EFFECTIVE NAME	CUSIP IDENTIFIER - HISTORICAL	EXCHANGE TICKER SYMBOL - HISTORICAL	COMPANY NAME - HISTORICAL
86079	665	19801204	19890131	11003510	BRST	BRISTOL GAMING CORP
86079	665	19890201	19920419	11003520	BRST	BRISTOL HOLDINGS INC
86079	665	19920420	19950822	84892010	SPTK	SPORTS TECH INC
86079	665	19950823	19970630	01662710	ALCM	ALL COMM MEDIA CORP
86079	665	19970701	20010729	57090710	MSGI	MARKETING SERVICES GROUP INC
86079	665	20010730	20011014	57090710	MKTG	MARKETING SERVICES GROUP INC
86079	665	20011015	20020331	57090720	MKTG	MARKETING SERVICES GROUP INC
86079	665	20020401	20030126	55308X10	MKTG	M K T G SERVICES INC
86079	665	20030127	20031127	55308X30	MKTG	M K T G SERVICES INC
86079	665	20031128	20040111	55308X30	MSGI	M K T G SERVICES INC
86079	665	20040112	20050208	58445910	MSGI	MEDIA SERVICES GROUP INC
86079	665	20050209	20061019	55357010	MSGI	M S G I SECURITY SOLUTIONS INC

1 PERMNO per stock   1 PERMCO per company

10 Different CUSIPs   5 Different Tickers   8 Different Names

wrds
The Global Standard for Business Research

Primary Identifiers

+ The Primary Identifiers are different
  + **Compustat**: GVKEY at firm level
  + **CRSP**: PERMNO (PERMCO) at security (firm) level

+ **PERMNO** is a unique five-digit permanent identifier assigned by CRSP to each security: *Permanent*, and *Unique (Not Recyclable)*.

+ **GVKEY** is a unique six-digit identifier and primary key for each company and assigned by Compustat: *Permanent*, and *Unique (Not Recyclable)*.

wrds

CRSP and Compustat CUSIP Match

Security level:

CRSP
MSFNAMES
(PERMNO)

→

CUSIP
=
SUBSTR(Cusip,1,8)

←

COMPUSTAT
NAMES
(GVKEY)
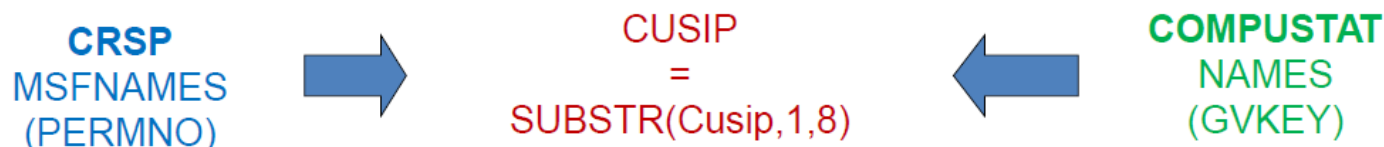
PERMNO-GVKEY match. It matches with the security that has price in Compustat.

Firm level:

CRSP
MSFNAMES
(PERMCO)

→

SUBSTR(**CUSIP**,1,6)
=
SUBSTR(**CUSIP**,1,6)

←

COMPUSTAT
NAMES
(GVKEY)
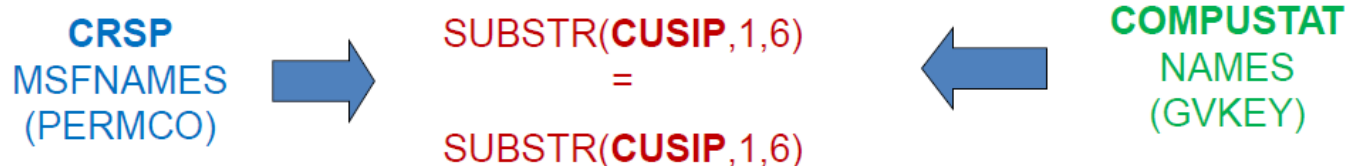
PERMCO-GVKEY match. It still needs to come up with the 'main' PERMNO for each PERMCO.

CRSP/Compustat Merged Data (CCM)

+ CCM database is to allow researchers to lookup and extract Compustat data using the CRSP identifiers.

+ CCM contains only Compustat data items, but can be searched by CRSP's PERMNO and PERMCO in addition to Compustat's GVKEY.

+ CCM maintains the matching table between CRSP and Compustat
  + Ideally, the matched firms are the ones listed in NYSE, AMEX, NASDAQ and ARCA and survived long enough to have the SEC filings

Advanced WRDS Features

Access Data Remotely on WRDS Server

+ What if a web query does not do all that you need it to do?

+ Web queries usually cannot handle the requests with > 2 conditions

+ Example: Find all companies in 1997 with sales greater than 1 billion, total assets greater than 5 billions, and with more than 30 years of publicly reported financial statements.

+ Two more ways to access finance data at WRDS

	Advantage	Disadvantage
UNIX	Fast remote data access	Batch Running
PC-SAS/Connect	Interactive Programming	Slow remote data access

+ Each method can perform a basic data extract and serve as the first step in a statistical analysis. The data can be analyzed remotely or imported into a local SAS system or even a second program to complete the analysis.

PC-SAS / Connect – Concept

+ SAS software installed on your Windows PC
  + Some local SAS dataset in your PC
  + Remote access WRDS data by connecting to WRDS server
    + Use WRDS powerful Unix server processing resources
    + Access Unix permanent (750MB) & temp (4.5TB) disk spaces

+ Steps:
  1. Connect to WRDS server
     ```
     %let wrds = wrds.wharton.upenn.edu 4016;
     options comamid = TCP remote = WRDS;
     signon username = _prompt_;
     ```
  2. Remote Submit
     ```
     rsubmit;
         {Program}
     endrsubmit;
     ```

PC-SAS / Connect – An Example

+ Example: Find all companies in 1997[①] with sales >= 1 billion[②], total assets >= 5 billions[③], and with >= 30 years of publicly reported financial statements[④].

```
/* PC-SAS/Connect Communication Block   */
%let wrds = wrds.wharton.upenn.edu 4016;
options comamid=TCP remote=WRDS;
signon username=_prompt_;

/* Submit SAS code to WRDS Unix Server */
rsubmit inheritlib=(work=pcwork);

/* SAS CODE submitted to WRDS Sever      */
proc sql;
create table pcwork.demo(
where= (fyear=1997 and sale>=1000 and at>=5000
and firm_age>=30)) as

select fyear, conm, tic, gvkey, sale, at,
(fyear – min(fyear)) as Firm_Age

from comp.funda where not missing(at) and
consol="C" and indfmt="INDL" and datafmt="STD"
and popsrc="D"

group by gvkey;
quit;
```

```
/* Remote SAS CODE submission Ends */
endrsubmit;

/* Sign off from PC-SAS/Connect    */
Signoff;
```

wrds

How to Access SAS Data in UNIX?
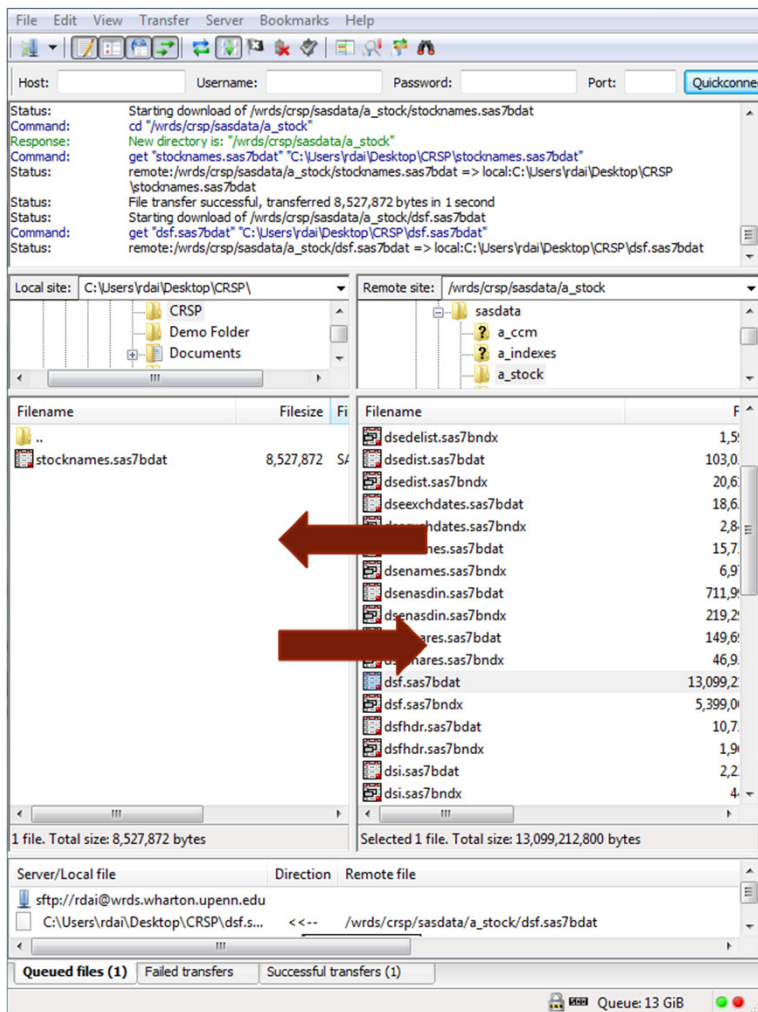
1. Install SFTP Client software in your PC (freeware such as FileZilla).

2. Set SFTP Client to be connected to the server: wrds.wharton.upenn.edu with your wrds (wharton) username and password.

3. Retrieve and Upload files in SAS and other formats among your local machine and remote UNIX server.

How to run SAS in UNIX?


```
wrds.wharton.upenn.edu - SecureCRT
File   Edit   View   Options   Transfer   Script   Tools   Help

✓ wrds.wharton.upenn.edu                                    ◁ ▷

Last login: Tue Nov  1 21:48:43 2011 from slc035.wharton.

*** Welcome to Wharton Research Data Services (WRDS) ***

WRDS is developed and run by the Wharton School at the Univer
sity of
Pennsylvania. We provide access to the databases that define
empirical,
seminal research in the fields of accounting, banking, econom
ics, finance,
insurance, management, marketing, public policy, risk managem
ent, and
statistics.

If you are not an authorized user of this system, disconnect
now.
All activity is strictly monitored. There may be severe crimi
nal and
civil penalties for unauthorized access or use of computing r
esources.

Please review the latest known issues of the interactive WRDS
 server at:

http://wrds-web.wharton.upenn.edu/wrds

Please report any problems to wrds-support@wharton.upenn.edu.

wrds(~)% sas demo.sas


Ready           ssh2: AES-256-CTR   29, 22   32 Rows, 61 Cols   VT100
```
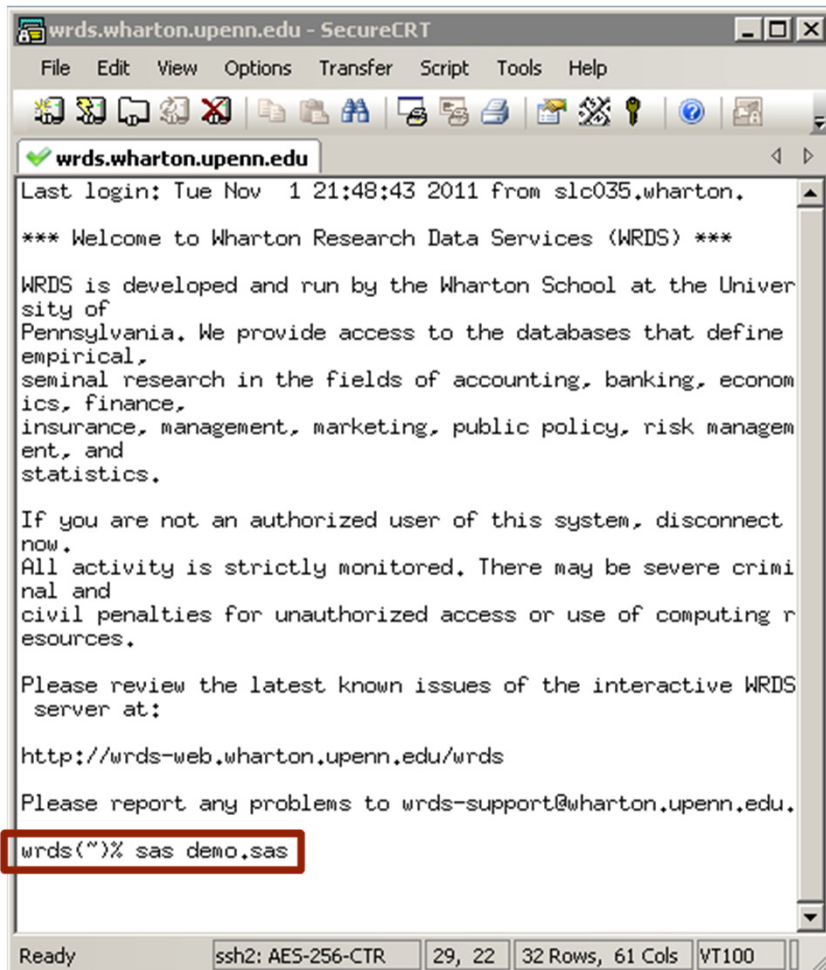
1. Install SSH Secure Shell (free) software in your PC.

2. Set SSH to be connected to the server: wrds.wharton.upenn.edu with your wrds (wharton) username and password.

3. Create, Edit and Run your SAS program using Unix commands.

wrds

I don't know SAS

+ WRDS support page has a lot of documentation and material for SAS and Unix.

+ There are a lot of useful SAS Books. Among them:

Research and Technical Support

+ Online Help
  + Database Manuals plus additional support documentation
    + [Data Overviews](#)
    + [Research Applications](#)
    + [Sample Programs](#)
    + [Variable Search](#)
    + [Company Search](#)
  + [WRDS Knowledge Base](#): FAQ archive of answers to common user questions

+ Email support at wrds-support@wharton.upenn.edu (Monday-Friday, 9a-5p EST)
  Researchers and Technical Experts ready to assist with:
  + Data extraction, merging, and management.
  + Programming and technical problems

Other Questions?