

# WRDS Quick Start Guide

---

## Table of Contents

What is WRDS? .....	3
Why should I use WRDS instead of other software in the lab? .....	3
Table 1: Finding the Right Database .....	3
How do you open WRDS? .....	4
Option 1: Through the webpage .....	4
Option 2: Through the icon .....	4
How do you create a WRDS account? .....	4
Option 1: Masters/PhD students/Research Assistant .....	4
Option 2: Undergraduate 24-hour pass.....	4
Available Databases in WRDS .....	4
Table 2: Database Descriptions .....	4
Downloading Data .....	12
FAQs.....	13

## What is WRDS?

WRDS stands for Wharton Research Data Services. It is a data delivery service covering a number of the databases to which we subscribe, such as CRSP and COMPUSTAT. A full list of available databases is provided later in this Quick Start Guide.

WRDS receives data directly from providers and posts the data to their system in SAS data sets (Statistical Analysis System).

## Why should I use WRDS instead of other available lab software?

The choice to use WRDS over other available software depends primarily on what type of data you are looking for. Please refer to the following table to determine which database is the appropriate platform for the information you are looking for:

**Table 1: Finding the Right Database**

	Real-Time Data	Historical Data	Bond Data	OTC Data	Dead Data	Equity Fundamentals	Real-time News	Excel Downloads	Real-Time Excel Link
Bloomberg	E	E	E	E	E	E	E	F	E
Capital IQ	M	E	N	N	M	EE	F	E	F
Datastream	N	EE	E	E	E	M	N	E	N
FactSet	E	E	E	NA	N	E	E	E	E
MorningStar Direct	E	F	F	NA	NA	E	E	E	NA

*EE=Fantastic, E = Excellent, F = Fair, M = Marginal, N = None*

You can find Quick Start Guides for the other software listed above in the Quick Start Guides folder on any lab computer desktop.

Other advantages to using WRDS include the ability to retrieve data over the internet, and the ability for data to be saved in a variety of formats, including SAS Windows\_32 Dataset, dBase file, STATA file, or SPSS file, in addition to common formats such as txt, xls or html.

## How do you open WRDS?

### Option 1: Through the webpage

Just follow this link to access WRDS through your internet browser:

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

### Option 2: Through the icon

Click on the WRDS icon (for visual see below) on the desktop of any lab computer to access the software



## How do you create a WRDS account?

### Option 1: Faculty/Masters students/PhD students/Research Assistants

If you are a Faculty member, Visiting Faculty, PHD student, Masters student or a Research Assistant at U of T, you already have access to WRDS. Please ask your supervisor to provide any further registration instruction.

### Option 2: Undergraduate 24-hour pass

If you are an undergraduate who needs to use WRDS for a course you may apply for a 24-hour pass using a lab computer. On the home-page of WRDS just enter in your UTOR-id email to apply for access. You may renew your 24-hour access anytime you wish.

## Available Databases in WRDS

The following table provides descriptions to all the databases that we currently have subscriptions to, and are available, in WRDS.

**Table 2: Database Descriptions**

Database	Introduction
AuditAnalytics	AuditAnalytics provides detailed audit information on over 10,000 accounting firms and 150,000 active audits. It tracks auditing details such as the auditor name and fees, as well as auditor reports on financial statements disclosed.
Bank Regulatory	The Bank Regulatory Database contains five databases for regulated depository financial institutions. These databases provide accounting data for bank holding companies, commercial banks, savings banks, and savings and loans institutions. The source of the data comes from the required regulatory forms filed for supervising purposes.

Blockholders	This dataset contains standardized data for blockholders of 1,913 companies. The data was cleaned from biases and mistakes usually observed in the standard source for this particular type of data. Blockholders' data is reported by firm for the period 1996-2001.
CBOE Indexes	The CBOE (Chicago Board Options Exchange) Volatility Index® (VIX®) is a key measure of market expectations of near-term volatility conveyed by S&P 500 stock index option prices. The VIX measures the market's expectation of 30-day volatility in a way that conforms to the latest thinking and research among industry practitioners. It is based on S&P 500 index option prices and incorporates information from the volatility "skew" by using a wider range of strike prices rather than just at-the-money series.
COMPUSTAT	A market database published by Standard and Poor's. The comprehensive Compustat database provides company data going back 40 to 50 years on over 65,000 securities, as of 2010. The type of information published by Compustat include: Global Industry Classification Standards (GICS), pricing data, earnings data, insider and institutional holdings, and other information directed at investors and analysts.
Center for Research in Security Prices (CRSP)	CRSP (Center for Research in Security Prices) maintains the most comprehensive collection of security price, return, and volume data for the NYSE, AMEX and NASDAQ stock markets. Additional CRSP files provide stock indices, beta- and cap-based portfolios, treasury bond and risk-free rates, mutual funds, and real estate data. - CRSP's unique historical Indices Databases serve as benchmarks for the investment community, and as a foundation for academic research. - The CRSP US Treasury and Inflation Series contain returns and index levels on the US Government Bond Fixed Term Index Series, and the Risk Free Rates File. - The CRSP/COMPUSTAT Merged Database allows for concurrent database access to CRSP's stock data and Compustat's fundamental data. The CRSP Link accurately maps complex, many-to-many relationships over time between CRSP's unique permanent identifiers (PERMNO and PERMCO), and Compustat's unique permanent identifier (GVKEY).
CUSIP	CUSIP provides CUSIP numbers, standardized descriptions and additional data attributes for over 5 million corporate, municipal and government securities offered in North America.
Direct Marketing Education Foundation (DMEF)	Contains four individual data sets, each containing customer buying history for about 100,000 customers of nationally known catalog and non-profit database marketing businesses.

Dow Jones	<p>The Dow Jones Averages are comprised of The Daily and Monthly Dow Jones Composite (DJA), as well as The Dow Jones Industrial (DJI), The Dow Jones Transportation (DJT), The Dow Jones Utility (DJU), The Dow 10, and The Dow 5. The Dow Jones Industrial Average, the best known U.S. stock index, and the Dow Jones Transportation Average, the oldest U.S. stock index were created by Charles Henry Dow, co-founder of Dow Jones &amp; Company, and maintained and reviewed by the editors of The Wall Street Journal. Dow Jones Indexes is the source of Dow Jones information and the provider of this data.</p> <p>The Dow Jones Total Return Indexes are introduced by Dow Jones Indexes as part of the Dow Jones Total Market Index Series. The Total Return Indexes account for reinvested dividends. Data for the Total Return Indexes is available after 1987.</p>
Eventus	<p>Eventus performs event studies using data read directly from CRSP stock databases or pre-extracted from any source. The Eventus system includes utility programs to convert calendar dates to CRSP trading day numbers, convert CUSIP identifiers to CRSP permanent identification numbers, and extract event study cumulative or compounded abnormal returns for cross-sectional analysis.</p>
FactSet	<p>Provides financial data and analysis for financial professionals, including investment managers, hedge funds and investment bankers. FactSet consolidates data on global markets, public and private companies, and equity and fixed-income portfolios.</p>

Fama-French & Liquidity Factors	<p>The Fama-French Portfolios are constructed from the intersections of two portfolios formed on size, as measured by market equity (ME), and three portfolios using the ratio of book equity to market equity (BE/ME) as a proxy for value. Returns from these portfolios are used to construct the Fama-French Factors. Eugene Fama and Kenneth French showed that their factors capture a statistically significant fraction of the variation in stock returns (see “Common Risk Factors in the Returns on Stocks and Bonds”, Journal of Financial Economics 33, 1993). The Fama-French data source is Kenneth French’s web site at Dartmouth.</p> <p>The Pastor-Stambaugh Liquidity series are described by L. Pastor and R. F. Stambaugh in “Liquidity risk and expected stock returns” (2003, Journal of Political Economy 111, 642-685). This set includes ‘non-traded’ and ‘traded’ liquidity factors, with the latter series derived from dividing common stocks (in the CRSP monthly stocks file data) into 10 groups based on each stock’s sensitivity to the ‘non-traded’ liquidity innovation factor (as described in the paper).</p> <p>The Sadka Liquidity measures are described in R. Sadka in “Momentum and Post-Earnings- Announcement Drift Anomalies: The Role of Liquidity Risk” (Journal of Financial Economics 80, 309-349). The measures are non-traded, market-wide, undiversifiable risk factors. Price impact is separated into permanent (variable) and transitory (fixed) price effects.</p>
Federal Reserve Bank	<p>The Federal Reserve Bank Reports contain three databases collected from Federal Reserve Banks: Foreign Exchange Rates (over 30 world currencies), Interest Rates, and FRB-Philadelphia State Indexes.</p>
I/B/E/S	<p>I/B/E/S International Inc. created their Academic Research Program over 30 years ago to provide both summary and individual analyst forecasts of company earnings, cash flows, and other important financial items, as well as buy-sell-hold recommendations. In 2000, I/B/E/S was integrated with Thomson Reuters / First Call.</p>
IHS Global Insight	<p>The "Global Insight Series Attributes" web query provides a means to save time in using the "Global Insight Series Data" web query. It also provides a good way to explore the breadth and depth of GI data coverage without processing the large volume of data values.</p> <p>The “Global Insight Series Data” web query provides access to over 1.5 million data series, from over 200 countries and international entities.</p>

The Institute for the Study of Security Markets (ISSM)	The Institute for the Study of Security Markets (ISSM) database contains tick-by-tick data covering the NYSE and AMEX between 1983 and 1992, and NASDAQ between 1987 and 1992. Each year of data is divided into two files, one for trades and one for quotes.
Markit	<p>- Market CDS Pricing: Provides CDS composite and contributor level data on approximately 2,800 individual entities. Applying a data cleansing process, Markit aggregates marks from sell-side contributors to generate its composite price. The data are provided on a daily basis starting in 2001 (with 810 reference entities) and are available by entity, tier, currency and restructuring clause.</p> <p>- RED (Reference Entity Database): a database of confirmed Reference Entities and their Reference Obligations for use in trading, documenting, and processing credit derivatives transactions. RED also allows the data user to track CDS reference entities through mergers, de-mergers and other corporate actions.</p>
MFLINKS	The MFLINKS tables provide a reliable means to link CRSP Mutual Fund (MFDB) data that covers mutual fund performance, expenses, and related information to equity holdings data in the Thomson Reuters Mutual Fund Ownership data (formerly known as the CDA S12 data). Using MFLINKS allows a researcher to gather detail on holdings either for particular funds or fund families and groups of funds at specific points in time.
Option Metrics	Ivy DB OptionMetrics is a comprehensive source of historical price and implied volatility data for the US equity and index options markets. It contains historical prices of options and their associated underlying instruments, correctly calculated implied volatilities, and option sensitivities. With Ivy DB OptionMetrics, you'll be able to backtest trading strategies, evaluate risk models, and perform research on all aspects of options investment.

OTC Markets	<p>OTC Markets Group's Open, Transparent and Connected financial marketplaces provide investors with the information necessary to intelligently analyze, value and trade 10,000 U.S. and global securities through the broker of their choice. To create clarity in the investment process, we organize these securities into three tiered marketplaces—OTCQX, OTCQB and OTC Pink—based on the quality and quantity of information the companies make available.</p> <p>Our marketplaces empower companies to provide the levels of transparency necessary to create more efficient markets. Together with our SEC regulated Alternative Trading System, OTC Link® ATS, the combined transparency and transactional efficiency creates a compelling trading experience for investors. OTC Markets Group (OTCQX: OTCM) is a public company traded on OTCQX under the symbol OTCM.</p>
Penn World Tables	<p>The Penn World Tables provides national income accounts-type of variables converted to international prices. The homogenization of national accounts to a common numeraire allows valid comparisons of income among countries.</p>
Philadelphia Stock Exchange (PHLX)	<p>The PHLX trades more than 2,800 stocks, 740 equity options, 12 sector index options and 100 currency pairs. It has created a unique market structure for the trading of currency options. Titled the United Currency Options Market (UCOM), it provides unprecedented flexibility and the availability of both standardized and customized currency options. Customized currency options offer choice of expiration date, strike (exercise) price, premium payment and any combination of 10 currencies currently available for a total of 100 possible currency pairs.</p>
SEC Order Execution	<p>On November 15, 2000, the SEC adopted new rules aimed at improving public disclosure of order execution and routing practices. As a result of Rule 11Ac1-5, market centers that trade national market system securities must make monthly, electronic disclosures of basic information concerning their quality of executions on a stock-by-stock basis, including how market orders of various sizes are executed relative to the public quotes and information about effective spreads - the spreads actually paid by investors whose orders are routed to a particular market center. In addition, market centers must disclose the extent to which they provide executions at prices better than the public quotes to investors using limit orders.</p>

Trade and Quoting (TAQ)	<p>Trade and Quote: Contains intraday transactions data (trades and quotes) for all securities listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), as well as Nasdaq National Market System (NMS) and SmallCap issues.</p> <p>Reg SHO - NYSE Short Sales: Contains data related to SEC Regulation SHO per the terms of the SRO requirement to make Short Sale transactions data available on a monthly basis for trade dates beginning January 2005 through the pilot period ending 2007. (SROs are self-regulatory organizations in the securities industry, such as the NYSE, that create and enforce rules for its members based on the federal securities laws.)</p> <p>Millisecond Trade and Quote: Trades and Bids are tracked at the millisecond level (HHMMSSXXX).</p>
Thomson Reuters	<p>Thomson Reuters is a global provider of data, analysis and information tools offering the most comprehensive range of indispensable, market-leading services and customizable solutions to help clients make better decisions, be more productive and achieve superior financial results.</p> <ul style="list-style-type: none"> <li>- Thomson-Reuters Institutional Holdings (13F) Database provides Institutional Common Stock Holdings and Transactions, as reported on Form 13F filed with the SEC. This database contains ownership information by institutional managers with \$100 million or more in Assets Under Management.</li> <li>- The Insider Filing Data Feed (IFDF) is designed to capture all U.S. insider activity as reported on Forms 3, 4, 5, and 144 in line-by-line detail. In addition, a unique cleansing process is used by Thomson Reuters that adds value by making systematic interpretations on the accuracy of as-reported data and inserts "cleansed" fields for comparison.</li> <li>- Thomson-Reuters Mutual Fund Holdings database provides security holding information for all registered mutual funds that report their holdings with the SEC, plus 3,000 global funds. Reported securities include all NYSE/AMEX/NASDAQ, Toronto, and Montreal common stocks.</li> <li>- WRDS-Thomson Reuters LPC provides users with access to Thomson Reuters LPC's robust database of detailed terms and conditions on over 240,000 loan transactions.</li> </ul>

TRACE	<p>Trade Reporting and Compliance Engine is FINRA's over-the-counter (OTC) corporate bond market real-time price dissemination service. The information collected and disseminated for all publicly traded corporate bonds by TRACE includes the time of execution, price, yield, and volume.</p>
WRDS SEC Analytics Suite	<p>WRDS SEC Analytics Suite with NEW SEC Sentiment and Readability data is positioned for broad business usage – from due diligence and forensic accounting to disclosure research and investment management.</p> <p>Simplify research—quickly and easily develop tailored datasets from all SEC filings, parsing millions of regulatory reports based on custom criteria. The ready-to-use parsed contents and filings archives are updated daily, allowing users to quickly explore data and effectively retrieve widely-used information including; the full list of exhibits, filing size, SEC acceptance date and time, conformed period of report, historical information on state of incorporation and headquarters, and SIC code.</p> <p>The WRDS SEC Readability and Sentiment database extends the SEC Analytics Suite by providing, for the first time to academic researchers, a clean set of readability and sentiment measures for every filing since 1994.</p>

---

Zacks

Zacks Investment Research, based in Chicago, Ill., has been a leading provider of research, market data, and quantitative models to institutional investment management firms in the US and Canada for over 30 years. Recognized for quality, consistency and reliability, Zacks provides institutional and individual investors with the analytical tools and financial information necessary to the success of their investment process. Founded in 1978, Zacks' early contribution to investment analysis was the discovery that earnings per share estimate revisions are the most powerful force affecting stock prices. This discovery is built into the Zacks Rank proprietary methodology for predicting stock price performance. The Zacks Rank has produced average annual returns in excess of 28% since 1988. The premier source of analysts' earnings forecasts, today Zacks produces data feeds for estimates, ratings, earnings report data, fundamental data, and institutional holdings for US and Canadian traded equities, as well as investment research reports and research software tools for investors.

Zacks data can be used to empirically analyze analysts' forecasts and their revisions, price targets and recommendations (e.g., construct popular measures such as consensus forecast, earnings surprises, analyst track records, etc).

---

*Note: subscription last updated Summer 2015*

## Downloading Data

First, you must choose a particular database. Once in your database of choice, you must pick your dataset. After choosing your dataset, WRDS will take you through a step-by-step process to downloading your data, also known as your 'query', in WRDS. We briefly show you the steps in this guide:

1. What date range do you want to use?
2. How would you like to search this dataset?
3. What variables do you want in your query?
  - This step is simply asking which data you would like (i.e. Company Name, Ticker Symbol, etc.)
4. How would you like the query output?
  - This step is just asking you in what format would you like your data (i.e. format for excel, STATA, etc.)
  - Formats available: .txt, .csv, .xlsx, .htm, .sas7bdat, .dbf, .dta, .sav

## FAQs

**Q:** What does it mean by ACCHG/ACT/ADD4/PERMCO/...? (any variables in abbreviation)

**A:** At the top of each query page you will find a link titled 'Variable Descriptions'. This will provide a list of all the abbreviations you might encounter in each WRDS database.

**Q:** I was downloading/pending a query but I happened to close the page. Where can I find my query data?

**A:** Your query won't be cancelled if you close the page. From the home screen, click the tab 'Recent Queries', and you should find your query.

**Q:** Do you provide any video training on WRDS?

**A:** Yes. If you click the 'Support' tab at the top right of your page, under 'Getting Started' you will see a link to 'Videos Tutorials & Presentations'. After following this link, you will find Beginner, Intermediate, and Advanced training videos. Furthermore, once you enter a particular database if you click on 'Manuals and Overviews' there are often video aids.

## Contact

Have more questions concerning WRDS? Contact this year's WRDS lab assistant specialist:

Laura Murphy

email: lc.murphy@mail.utoronto.ca