

## USING FTP TO SEND OR GET DATA TO/FROM AN ANONYMOUS SITE FROM OUR UNIX<sup>1</sup>

Many sites make data and other files available for your use. For example, the 2000 Census data and other Census Bureau data product are available at the following FTP address: [www2.census.gov](http://www2.census.gov). You may use FTP to get the data. For anonymous FTP sites, you must know the FTP address. For the name, enter “anonymous.” For the password, the site usually requests that you enter your email address. For example, I would enter “apchelp@albany.edu.”

You will note that some ftp commands are the same as used in UNIX. For example, “dir” is used to find the contents of a directory, and “cd” is used to change directory. And, the results of “dir” look similar to those found when using UNIX. In the example presented later, you will see the following results from “dir”. As you would see in a UNIX session, the “d” in the first column indicates a directory.

```
drwxr-xr-x  8      65603      0      4096 Sep 11      2003 census_1980
drwxr-xr-x  95     65603      0      4096 May 10      2005 census_1990
drwxr-xr-x  3       0          0      4096 Sep 03      2003 census_2000
```

Sometimes data are stored in compressed format. A compressed file will have the extension of “.gz” if it was compressed with the utility “gzip”, “.Z” if it was compressed with “compress” and “.zip” if it is a file (or files) that was created with “pkzip”. It is possible to uncompress .gz and .Z files in the process of FTPing the file, but I do not recommend this. It will take longer to transport the file, and it will take more room to store. Files created with pkzip can be uncompressed on a personal computer with “pkunzip.exe”, or, in some circumstances, with the utility “unzip” on the UNIX. All compressed files must be transported as binary files.

The default for FTP is ASCII, so, if the file to be transported is binary, you must issue the command “binary”.

Following is an example of an FTP session on the UNIX in which I retrieve a binary file. My commands are in bold. Before you start, you will want to decide where you will want the file to go. It must go to an existing directory to which you have write access, and which has adequate room for the file. You can change to that directory before you begin your FTP session, or use the “lcd” command as demonstrated below to change the local directory from within the FTP session. Commands in FTP sessions (as with UNIX sessions) are case sensitive so use upper and lower case appropriately.

```
liska% ftp www2.census.gov
Connected to ftp2.census.gov.
220 You have accessed ftp2.census.gov, a Government computer!
Name (www2.census.gov:apchelp): anonymous
331 Please specify the password.
Password: myid@mynode (Use your own user id and node.)
230 Login successful.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> dir
```

---

<sup>1</sup> Originally prepared by Patty Glynn, University of Washington in March 1998 and modified by Hui-shien Tsao, University at Albany, March 10, 2006.

drwxrwsr-x	5	0	8139	4096	Jun 14	2005	AHS
drwxr-s---	4	0	8115	4096	May 19	2005	acdc
drwxrwsr-x	3	8053	5005	4096	Aug 30	2005	acs
drwxrwsr-x	5	8053	5114	4096	Aug 13	2004	acs2002
drwxrwsr-x	5	8105	5116	4096	Oct 06	2004	acs2003
drwxrwxr-x	5	66412	5005	4096	2/2/2006	10:40	acs2004
drwxrwsr-x	3	0	5115	4096	Oct 06	2004	acs_special_tabs
drwxr-xr-x	8	65603	0	4096	Sep 11	2003	census_1980
drwxr-xr-x	95	65603	0	4096	May 10	2005	census_1990
drwxr-xr-x	3	0	0	4096	Sep 03	2003	census_2000
drwxr-xr-x	5	0	0	4096	Mar 30	2005	dadso
drwx-----	2	1359	0	4096	Dec 03	2004	dbadrak
drwxrwsr-x	8	1360	1001	4096	12/6/2006	13:01	econ2002
drwxrwsr-x	4	0	1001	4096	10/18/2006	16:35	econ2003
drwxrwsr-x	4	0	1001	4096	12/28/2006	16:39	econ2004
drwxrws---	3	0	8119	4096	2/22/2006	11:58	evaluate
-rw-r--r--	1	0	0	318	May 15	2001	favicon.ico
drwxr-x---	4	65603	48	4096	12/8/2006	14:40	ftp
drwxrwsr-x	9	1367	7000	4096	1/25/2006	11:24	geo
drwxrwsr-x	22	0	4002	4096	May 26	2005	govs
drwxrwsr-x	3	0	8136	4096	May 10	2005	hhes
lrwxrwxrwx	1	0	0	12	May 07	2005	inc -> pub/incoming
							incoming ->
lrwxrwxrwx	1	0	0	12	May 07	2005	pub/incoming
drwxrwsr-x	48	0	7008	4096	1/25/2006	11:27	lehd
drwx-----	2	0	0	4096	11/21/2006	13:32	lost+found
							outgoing ->
lrwxrwxrwx	1	0	0	12	May 07	2005	pub/outgoing
drwxr-sr-x	5	0	7000	4096	Jul 11	2003	plmap
drwxrwsr-x	2	0	4006	4096	Nov 20	2003	pop
drwxr-xr-x	7	1386	0	4096	Jun 29	2005	prod2
drwxr-xr-x	4	0	0	4096	Mar 17	2005	pub
drwxrws---	3	0	2698	4096	Jun 16	2005	restricted
-rw-r--r--	1	0	0	0	Nov 19	2001	robots.txt
drwxr-xr-x	2	65603	0	4096	Aug 29	2005	tms_data
dr-xr-x---	2	0	48	8192	Jun 14	2004	web

226 Directory send OK.

2249 bytes received in 0.01 seconds (212.51 Kbytes/s)

ftp> **cd census\_2000**

250 Directory successfully changed.

ftp> **dir**

200 PORT command successful. Consider using PASV.

150 Here comes the directory listing.

drwxrwsr-x	22	1360	1001	4096	12/1/2006	13:56	datasets
------------	----	------	------	------	-----------	-------	----------

226 Directory send OK.

66 bytes received in 0.0088 seconds (7.30 Kbytes/s)

ftp> **cd datasets**

250 Directory successfully changed.

ftp> **dir**

200 PORT command successful. Consider using PASV.

150 Here comes the directory listing.

drwxr-xr-x	56	1373	2698	4096	Jun 07	2002	100_and_sample_profile
drwxr-xr-x	4	0	0	4096	Mar 13	2003	108_Congressional_Districts
drwxr-xr-x	4	1360	7002	4096	Jan 06	2005	109_Congressional_Districts
drwxrwxr-x	3	1360	7002	4096	Oct 01	2003	AIAN_SF
drwxrwxr-x	2	1388	7002	4096	Oct 01	2003	CQS
lrwxrwxrwx	1	0	0	7	May 06	2005	EEO_2000 -> eeo2000
drwxr-xr-x	6	1360	7005	4096	Apr 30	2003	Island_Areas
drwxr-xr-x	4	0	0	4096	Aug 06	2003	PUMS
drwxr-xr-x	2	1388	0	4096	Jun 09	2004	Race_Hispanic_Latino_Summary_File
drwxr-xr-x	2	1388	7002	4096	May 05	2003	Sim_Hispanic_Totals
drwxr-xr-x	56	1360	7002	4096	Oct 19	2002	Summary_File_1
drwxr-xr-x	54	1360	2698	4096	Jun 11	2003	Summary_File_1_Supplement
drwxr-xr-x	56	1360	7005	4096	Jan 25	2003	Summary_File_2
drwxr-xr-x	54	0	0	4096	Mar 27	2003	Summary_File_2_Supplement
drwxr-xr-x	55	1360	7005	4096	Sep 24	2002	Summary_File_3
drwxr-xr-x	55	0	0	4096	Jul 22	2003	Summary_File_4
drwxrwsr-x	6	14	1360	4096	12/8/2006 15:38		Summary_File_Extracts
drwxr-xr-x	57	0	0	4096	Apr 19	2002	demographic_profile
drwxrwsr-x	2	0	1001	4096	Dec 14	2004	determination
drwxrwsr-x	4	1400	1400	4096	Feb 24	2004	eeo2000
drwxrwxr-x	56	0	1001	4096	Oct 12	2001	redistricting_file--pl_94-171

226 Directory send OK.

1586 bytes received in 0.0078 seconds (198.82 Kbytes/s)

ftp> **cd Summary\_File\_1**

250 Directory successfully changed.

ftp> **dir**

200 PORT command successful. Consider using PASV.

150 Here comes the directory listing.

-rw-r--r--	1	1360	2698	23552	Aug 15	2001	0File_Structure_README.doc
-rw-r--r--	1	1360	2698	374	Aug 15	2001	0File_Structure_README.txt
-rw-r--r--	1	1360	2698	6004	Aug 15	2001	0File_Structure_README.wpd
drwxr-xr-x	2	1360	7002	4096	Oct 23	2002	0Final_National
-rw-r--r--	1	1360	2698	76800	Feb 05	2002	0README_SF1.doc
-rw-r--r--	1	1360	2698	11651	Feb 05	2002	0README_SF1.txt
-rw-r--r--	1	1360	2698	64023	Feb 05	2002	0README_SF1.wpd
drwxr-xr-x	2	1360	7001	4096	Nov 19	2001	Advance_National
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Alabama
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Alaska
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Arizona
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Arkansas
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	California
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Colorado

drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Connecticut
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Delaware
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	District_of_Columbia
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Florida
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Georgia
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Hawaii
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Idaho
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Illinois
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Indiana
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Iowa
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Kansas
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Kentucky
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Louisiana
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Maine
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Maryland
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Massachusetts
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Michigan
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Minnesota
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Mississippi
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Missouri
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Montana
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Nebraska
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Nevada
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	New_Hampshire
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	New_Jersey
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	New_Mexico
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	New_York
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	North_Carolina
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	North_Dakota
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Ohio
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Oklahoma
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Oregon
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Pennsylvania
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Puerto_Rico
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Rhode_Island
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	South_Carolina
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	South_Dakota
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Tennessee
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Texas
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Utah
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Vermont
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Virginia
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Washington
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	West_Virginia
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Wisconsin
drwxr-xr-x	2	1360	7001	4096	Aug 30	2001	Wyoming

226 Directory send OK.

4087 bytes received in 0.022 seconds (178.61 Kbytes/s)

ftp> lcd

```
Local directory now /home2/generic/apchelp
ftp> lcd /tmp
Local directory now /tmp
ftp> binary
200 Switching to Binary mode.
ftp> get 0README_SF1.txt
200 PORT command successful. Consider using PASV.
150 Opening BINARY mode data connection for 0README_SF1.txt (11651 bytes).
226 File send OK.
local: 0README_SF1.txt remote: 0README_SF1.txt
11651 bytes received in 0.053 seconds (212.87 Kbytes/s)
ftp> quit
221 Goodbye.
```

After completing the transfer, you should check to make sure that the transferred file is the same size AS the original.

```
malthus% dir
 12 -rw-----  1 apchelp  generic   11651 Mar 10 10:35 0README_SF1.txt
```

If you were sending a file instead of receiving a file, you would substitute “put” for “get”. The command would be something like: put /full/path/local/filename /full/path/remote/filename Often anonymous FTP sites have a directory called “incoming” where incoming files can be written.

It is also possible to use the “wild card” symbol (an asterisk) to get (type mget for “multiple get”) or send (type mput for “multiple put”) multiple files. For example I could have typed “mget chap\*.doc” to get all files that began with “chap” and ended with “.doc”.