

Data Analysis and Statistics

Jayant Murthy

The Indian Institute of Astrophysics

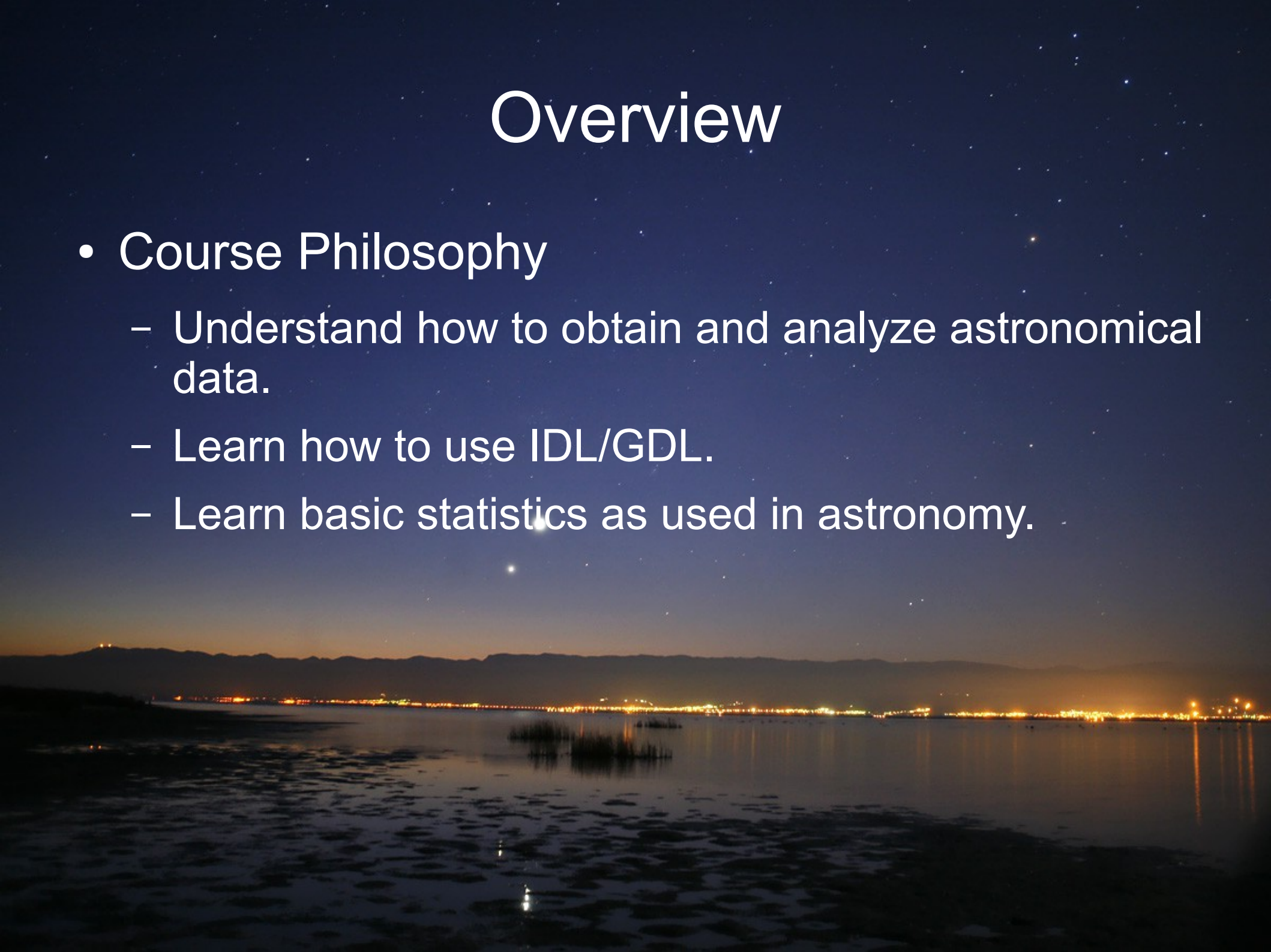
www.iiap.res.in

jmurthy@yahoo.com



Overview

- Course Philosophy
 - Understand how to obtain and analyze astronomical data.
 - Learn how to use IDL/GDL.
 - Learn basic statistics as used in astronomy.



Why GDL

- Because I know IDL.
 - GDL: free version of IDL.
 - Several Mac/Linux versions.
 - Extensive library of routines intended for astronomical use.
 - Easy prototyping for data analysis.
- Python may be preferred but I don't know it.

How to Install GDL

- <http://gnudatalanguage.sourceforge.net/index.php>
 - Can compile from source.
 - Binaries available from:
<http://gnudatalanguage.sourceforge.net/downloads.php>
- MACOS
 - I use MacPorts (<https://www.macports.org/>)
 - The name of the port is **gnudatalanguage** and the command once MacPorts is installed is *port install gnudatalanguage*.
- Linux
 - I have used the Debian version.

How to install under Windows

- I have installed VirtualBox, then installed Debian under VirtualBox and then GDL inside Debian.
 - Note that I don't now have access to Windows.
- Obtain VirtualBox from
 - <https://www.virtualbox.org/wiki/Downloads>
 - Go through the installation procedure. (It should be as easy as double-clicking on VirtualBox.exe.)
- Get Debian from <http://www.debian.org/distrib/netinst>
 - I used the “small files” which are ~280 MB in size but download more

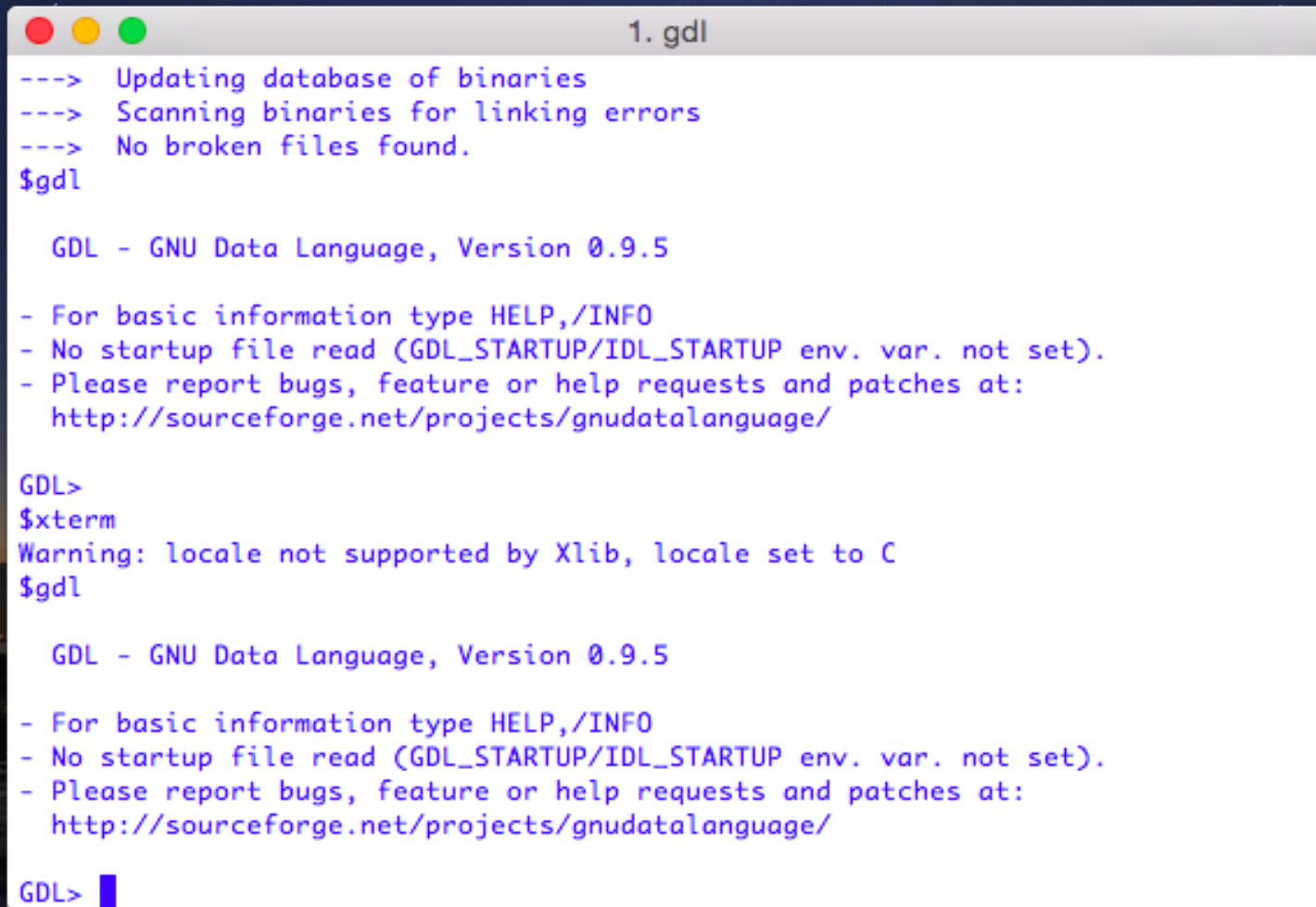
Fawlty IDL

- Fawlty IDL is available from <https://www.dropbox.com/sh/0vqq3pmu57rd04q/AADr8wWgx-7fIK76VKAL3IW7a?dl=0>
- I have not used it but there are a number of executables including Windows.



Once GDL is Installed

- Start by typing gdl from a terminal window.

A screenshot of a terminal window titled "1. gdl". The window shows the execution of the "gdl" command. It displays status messages about updating the binary database and scanning for errors. Then it shows the GDL version (0.9.5) and a list of instructions for users, including where to report bugs. The prompt changes from "\$gdl" to "GDL>". The user then types "\$xterm", which results in a warning about the locale and the prompt returning to "\$gdl". Finally, the GDL version and instructions are shown again, and the prompt returns to "GDL>".

```
1. gdl
---> Updating database of binaries
---> Scanning binaries for linking errors
---> No broken files found.
$gdl

GDL - GNU Data Language, Version 0.9.5

- For basic information type HELP,/INFO
- No startup file read (GDL_STARTUP/IDL_STARTUP env. var. not set).
- Please report bugs, feature or help requests and patches at:
  http://sourceforge.net/projects/gnudatalanguage/

GDL>
$xterm
Warning: locale not supported by Xlib, locale set to C
$gdl

GDL - GNU Data Language, Version 0.9.5

- For basic information type HELP,/INFO
- No startup file read (GDL_STARTUP/IDL_STARTUP env. var. not set).
- Please report bugs, feature or help requests and patches at:
  http://sourceforge.net/projects/gnudatalanguage/

GDL> 
```

GDL Basics

- Interpreted language.
 - Interactive commands have to be on a single line.
 - Loops can only be implemented in programs.
 - *Cheats shown later.*
 - Variables defined on the fly – no declaration.
 - *May lead to problems.*
 - Integrated graphics using X-Windows.

Calculator

```
1. gdl
Last login: Tue Jul  7 06:21:46 on ttys000
$gdl

GDL - GNU Data Language, Version 0.9.5

- For basic information type HELP,/INFO
- No startup file read (GDL_STARTUP/IDL_STARTUP env. var. not set).
- Please report bugs, feature or help requests and patches at:
  http://sourceforge.net/projects/gnudatalanguage/

GDL> print,5+5
      10
GDL> print,5*5
      25
GDL> print,5/5
       1
GDL> print,5-5
       0
GDL> print,5+5*5
      30
GDL> print,(5+5)*5
      50
GDL> print,5^2
      25
GDL> 
```

Variable Type

- Integer

```
1. gdl
- For basic information type HELP,/INFO
- No startup file read (GDL_STARTUP/IDL_STARTUP env. var. not set).
- Please report bugs, feature or help requests and patches at:
  http://sourceforge.net/projects/gnudatalanguage/

GDL> print,6-5
      1
GDL> print,6/5
      1
GDL> print,4/5
      0
GDL> print,32767
    32767
GDL> print,32767+1
   -32768
GDL> print,5-6
     -1
GDL> print,-32767
   -32767
GDL> print,-32767-1
   -32768
GDL> print,-32767-2
    32767
GDL> 
```


Other Types

- float, long, double, byte

```
1. gdl
$gdl

GDL - GNU Data Language, Version 0.9.5

- For basic information type HELP,/INFO
- No startup file read (GDL_STARTUP/IDL_STARTUP env. var. not set).
- Please report bugs, feature or help requests and patches at:
  http://sourceforge.net/projects/gnudatalanguage/

GDL> print,6./5
      1.20000
GDL> print,long(32767)+1
      32768
GDL> print,float(6)/5
      1.20000
GDL> print,long(6)/5
           1
GDL> print,1000*1000
      16960
GDL> print,10001*1000
     1000000
GDL> print,1000.*1000
     1.00000e+06
GDL> 
```

Variable Assignments

• =

```
1. gdl
GDL - GNU Data Language, Version 0.9.5

- For basic information type HELP,/INFO
- No startup file read (GDL_STARTUP/IDL_STARTUP env. var. not set).
- Please report bugs, feature or help requests and patches at:
  http://sourceforge.net/projects/gnudatalanguage/

GDL> a=5
GDL> b=5.
GDL> c=5l
GDL> d=double(5)
GDL> e=byte(5)
GDL> help
% At $MAIN$
A          INT          =          5
B          FLOAT        =      5.00000
C          LONG         =          5
D          DOUBLE       =      5.0000000
E          BYTE         =      5
Compiled Procedures:
$MAIN$

Compiled Functions:

GDL> █
```


A simple program

- vi first_program.pro
 - i (to insert)

```
data/          old/
$cd user/education/course/data_analysis/
$clear
a = 5.0
$pwd
/Users/jayanth/user/education/course/data_analysis
$ls
Lecture_1_Introduction.odp    figures
$mkdir programs
$cd programs
$vi first_program.pro
$more first_program.pro
a = 5.0
b = 10.0
print,a*b
for i = 0,10 do begin
    b = b*a
    print,b
endfor
end
$gdl

GDL - GNU Data Language, Version 0.9.5
- For basic information type HELP,/INFO
```

```
end
$gdl

GDL - GNU Data Language, Version 0.9.5
- For basic information type HELP,/INFO
- No startup file read (GDL_STARTUP/IDL_STARTUP env. var. not set).
- Please report bugs, feature or help requests and patches at:
  http://sourceforge.net/projects/gnudatalanguage/

GDL> .run first_program
% Compiled module: $MAIN$.
50.0000
50.0000
250.000
1250.00
6250.00
31250.0
156250.
781250.
3.90625e+06
1.95312e+07
9.76562e+07
4.88281e+08
GDL> █
```

Arrays

- Findgen
 - Fills with (float) integers
- Index starts with 0
- Row fills first then column.

```
2. gdl
GDL> a=findgen(10)
GDL> b=a*2
GDL> print,b
    0.00000    2.00000    4.00000    6.00000    8.00000   10.0000
   12.0000   14.0000   16.0000   18.0000
GDL> for i=0,9 do print,i,b(i)
    0    0.00000
    1    2.00000
    2    4.00000
    3    6.00000
    4    8.00000
    5   10.0000
    6   12.0000
    7   14.0000
    8   16.0000
    9   18.0000
GDL> a=findgen(2,2)
GDL> print,a*2
    0.00000    2.00000
    4.00000    6.00000
GDL> for i=0,1 do for j=0,1 do print,i,j,a(i,j)*2
    0    0    0.00000
    0    1    4.00000
    1    0    2.00000
    1    1    6.00000
```

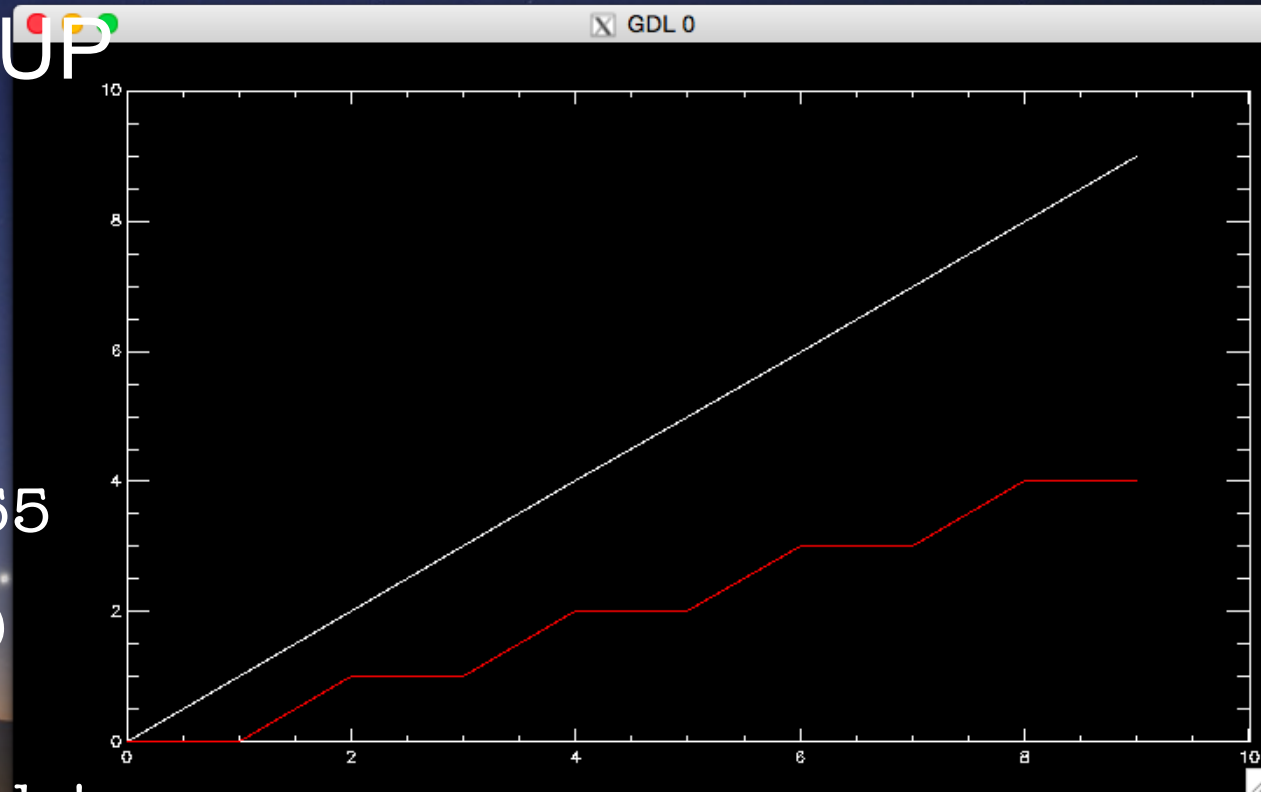

System Variables

- Begin with !
- Useful variables:
 - !PATH
 - !PI
 - !RADEG

```
2. gdl
!PI          FLOAT      =      3.14159
!DTOR        FLOAT      =      0.0174533
!RADEG       FLOAT      =      57.2958
!CONST       STRUCT     = -> !CONST Array[1]
!X           STRUCT     = -> !AXIS Array[1]
!Y           STRUCT     = -> !AXIS Array[1]
!Z           STRUCT     = -> !AXIS Array[1]
!VERSION     STRUCT     = -> !VERSION Array[1]
!MOUSE       STRUCT     = -> !MOUSE Array[1]
!ERROR_STATE STRUCT     = -> !ERROR_STATE Array[1]
!ERROR       LONG       =      0
!ERR         LONG       =      0
!ERR_STRING  STRING     = ''
!VALUES      STRUCT     = -> !VALUES Array[1]
!JOURNAL     LONG       =      0
!EXCEPT    INT        =      1
!MAP         STRUCT     = -> !MAP Array[1]
!CPU         STRUCT     = -> !CPU Array[1]
!DIR         STRING     = '/opt/local'
!GSHHS_DATA_DIR STRING   = ''
!STIME       STRING     = ''
!WARN        STRUCT     = -> !WARN Array[1]
!USERSYM     STRUCT     = -> !USERSYM Array[1]
GDL> !prompt="NO_PYTHON_HERE"
NO_PYTHON_HERE
```

Miscellaneous Useful Stuff

- Set GDL_STARTUP
- journal saves commands.
- GDL> journal
- GDL> defsysv,"!re",255
- GDL> plot,indgen(10)
- GDL> oplot,indgen(10)/2,col=!re



Library Programs

- <http://idlastro.gsfc.nasa.gov>
- Links to other resources
 - http://idlastro.gsfc.nasa.gov/other_url.html
- Adding libraries:
 - GDL_PATH variable tells GDL where to look.
 - In .bash_profile
 - `export`
`GDL_PATH="/Users/jayanth/user/idluser/idllib/pro:/opt/local/share/gnu`
`udatalanguage/lib:/Users/jayanth/user/data/Schlegel/CodeIDL"`