Understanding raw data:

- **Data sources**: Location where data originates
  - Report from a computer system
  - Selection from a large online database
  - Data table that has been manually entered
- **Data formats**
  - Tabular
  - XML
  - CSV
  - Excel
  - DB
  - JSON
- **Data types**
  - First-party data
    Data that was gathered from inside your own organization.
  - Second-party data
    Data that was gathered outside your organization but directly from the original source.
  - Third-party data
    Data gathered outside your organization and aggregated.
  - Geographic
  - Demographic
  - Time-based
  - Financial
  - Qualitative

Datetime functions

Code	Format	Example
%a	Abbreviated workday	Sun
%A	Weekday	Sunday
%b	Abbreviated month	Jan
%B	Month name	January
%c	Date and time	Sun Jan 1 00:00:00 2021

%d	Day (leading zeros)	01 to 31
%H	24 hours	00 to 23
%I	12 hours	01 to 12
%j	Day of year	001 to 366
%m	Month	01 to 12
%M	Minute	00 to 59
%p	AM or PM	AM/PM
%S	Seconds	00 to 61
%U	Week number (Sun)	00 to 53
%W	Week number (Mon)	00 to 53
%w	Weekday	0 to 6
%x	Locale's appropriate date representation	08/16/88 (None); 08/16/1988 (en_US); 16.08.1988 (de_DE)
%X	A locale's appropriate time representation	21:30:00 (en_US); 21:30:00 (de_DE)
%y	Year without century	00 to 99
%Y	Year	2022
%z	Offset	+0900
%Z	Time zone	EDT/JST/WET etc (GMT)

Code	Input Type	Input Example	Output Type	Output Example
`datetime.strptime("25/11/2022", "%d/%m/%Y")`	string	"25/11/2022"	DateTime	"2022-11-25 00:00:00"
`datetime.strftime(dt_object, "%d/%m/%Y")`	DateTime	"2022-11-25 00:00:00"	string	"25/11/2022"
`dt_object = datetime.strptime("25/11/2022", "%d/%m/%Y")` `datetime.timestamp(dt_object)`	string	"25/11/2022"	float (UTC timestamp in seconds)	1617836400.0
`datetime.strptime("25/11/2022", "%d/%m/%Y").strftime("%Y-%m-%d")`	string	"25/11/2022"	string	"2022-11-25"
`datetime.fromtimestamp(1617836400.0)`	float (UTC timestamp in seconds)	1617836400.0	DateTime	datetime.datetime(2021, 4, 7, 23, 0)
`datetime.fromtimestamp(1617836400.0).strftime("%d/%m/%Y")`	float (UTC timestamp in seconds)	1617836400.0	string	"'07/04/2021'"

`from pytz import timezone`
`ny_time = datetime.strptime("25-11-2022  09:34:00-0700", "%d-%m-%Y %H:%M:%S%z")`
`Tokyo_time = ny_time.astimezone(timezone('Asia/Tokyo'))`
	string	NewYork timezone "25-11-2022 09:34:00-0700"	DateTime	Tokyo timezone 2022, 11, 26, 1, 34, JST+9:00:00 STD>

`datetime.strptime("20:00", "%H:%M").strftime("%I:%M %p")`	string	"20:00"	string	"08:00 PM"
`datetime.strptime("08:00 PM", "%I:%M  %p").strftime("%H:%M")`	string	"08:00 PM"	string	"20:00"