

Pandas DataFrames

The official project homepage

- Goal
 - Extend what we learned about Series objects in the previous tutorial to their 2D counterpart - DataFrames
 - Develop some tools for dealing with missing data (not exhaustive, but a start)

DataFrames

Pandas quick start guide for DataFrames

- A DataFrame (DF) is a labeled data struture that can be thought of as a 2D extension of the Series objects that we discussed in the first part of the tutorial
- A DF can accept many types of input, multiple Series, a dict of 1D arrays, another DF, etc
- Like a Series, DFs contain data values and their labels. Because we're now dealing with a 2D structure, we call the **row labels the index argument** and the **column labels the column argument**.
 - Like a Series, if you don't explicitly assign row and column labels, then they will be auto-generated (but not as useful as specifying the labels yourself!)

Much of what we learned about Series objects will generalize to DFs, so here we'll focus on some of key functionality that might not be obvious based on the first part of the tutorial.

One more quick note: if using an older version of Python (earlier than 3.6) and Pandas (earlier than 0.23) and you create a DF from a dict without explicitly specifying column names, then the column names will be entered into the DF based on lexical order.

Import libs

```
In [1]: # import a pandas object and also a few specific functions that we'll use
import pandas as pd
from google.colab import files
```

Upload a file to the /content folder on google colab

- Select the file you want to upload (from the csv file that I sent out)
- It will load into your 'contents' folder
- Then you can interact with it just like a normal file on your hardrive

```
In [ ]: !ls
annual_temp2.csv  sample_data/

In [ ]: files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please run this cell to enable.

Saving annual_temp.csv2 to annual_temp.csv

```
Out[ ]: ['annual_temp.csv2.csv', b'Source,Year,Mean\nvngcag,2016,0.9363\nvngistemp,2016,\n\nvngcag,2015,\n\nvngistemp,2015,0.87\n\nvngcag,2014,0.7408\n\nvngistemp,2014,0.624\n\nvngcag,2013,0.624\n\nvngistemp,2013,0.63\n\nvngcag,2011,0.5788\n\nvngistemp,2011,0.6\n\nvngcag,2010,0.6367\n\nvngistemp,2009,0.64\n\nvngcag,2008,0.5417\n\nvngistemp,2008,0.54\n\nvngcag,2007,0.61\n\nvngistemp,2007,0.66\n\nvngcag,2006,0.6125\n\nvngistemp,2006,0.63\n\nvngcag,2005,0.6585\n\nvngistemp,2005,0.62\n\nvngcag,2004,0.6203\n\nvngistemp,2004,0.55\n\nvngcag,2003,0.63\n\nvngistemp,2003,0.62\n\nvngcag,2002,0.6026\n\nvngistemp,2002,0.63\n\nvngcag,2001,0.5473\n\nvngistemp,2001,0.55\n\nvngcag,2000,0.4262\n\nvngistemp,2000,0.42\n\nvngcag,1999,0.4438\n\nvngistemp,1999,0.42\n\nvngcag,1998,0.4344\n\nvngistemp,1998,0.44\n\nvngcag,1997,0.4587\n\nvngistemp,1997,0.48\n\nvngcag,1996,0.3228\n\nvngistemp,1996,0.35\n\nvngcag,1995,0.4377\n\nvngistemp,1995,0.46\n\nvngcag,1994,0.3409\n\nvngistemp,1994,0.32\n\nvngcag,1993,0.2853\n\nvngistemp,1993,0.24\n\nvngcag,1992,0.294\n\nvngistemp,1992,0.23\n\nvngcag,1991,0.4055\n\nvngistemp,1991,0.43\n\nvngcag,1990,0.493\n\nvngistemp,1990,0.41\n\nvngcag,1989,0.4087\n\nvngistemp,1989,0.39\n\nvngcag,1988,0.41\n\nvngcag,1987,0.3696\n\nvngistemp,1987,0.33\n\nvngcag,1986,0.2296\n\nvngistemp,1986,0.19\n\nvngcag,1985,0.1342\n\nvngistemp,1985,0.12\n\nvngcag,1984,0.149\n\nvngcag,1983,0.15\n\nvngcag,1982,0.14\n\nvngcag,1981,0.1341\n\nvngistemp,1981,0.09\n\nvngcag,1980,0.1354\n\nvngistemp,1980,0.13\n\nvngcag,1979,0.2273\n\nvngistemp,1979,0.17\n\nvngcag,1978,0.1123\n\nvngistemp,1978,0.07\n\nvngcag,1977,0.15\n\nvngcag,1976,0.18\n\nvngcag,1975,0.0792\n\nvngistemp,1975,0.11\n\nvngcag,1974,0.0343\n\nvngistemp,1974,0.02\n\nvngcag,1973,0.0264\n\nvngistemp,1973,0.01\n\nvngcag,1972,0.0264\n\nvngistemp,1972,0.01\n\nvngcag,1971,0.0313\n\nvngistemp,1971,0.03\n\nvngcag,1970,0.0372\n\nvngistemp,1970,0.02\n\nvngcag,1969,0.0329\n\nvngistemp,1969,0.07\n\nvngcag,1968,0.0286\n\nvngistemp,1968,0.07\n\nvngcag,1967,0.0231\n\nvngistemp,1967,0.02\n\nvngcag,1966,0.0227\n\nvngistemp,1966,0.05\n\nvngcag,1965,0.0278\n\nvngistemp,1965,0.1\n\nvngcag,1964,0.1489\n\nvngistemp,1964,0.2\n\nvngcag,1963,0.1088\n\nvngistemp,1963,0.06\n\nvngcag,1962,0.0883\n\nvngistemp,1962,0.03\n\nvngcag,1961,0.075\n\nvngistemp,1961,0.05\n\nvngcag,1960,0.0204\n\nvngistemp,1960,0.02\n\nvngcag,1959,0.0596\n\nvngistemp,1959,0.03\n\nvngcag,1958,0.1095\n\nvngistemp,1958,0.07\n\nvngcag,1957,0.0488\n\nvngistemp,1957,0.03\n\nvngcag,1956,0.0284\n\nvngistemp,1956,0.02\n\nvngcag,1955,0.1354\n\nvngistemp,1955,0.15\n\nvngcag,1954,0.1165\n\nvngistemp,1954,0.13\n\nvngcag,1953,0.0952\n\nvngistemp,1953,0.08\n\nvngcag,1952,0.0248\n\nvngistemp,1952,0.01\n\nvngcag,1951,0.0132\n\nvngistemp,1951,0.01\n\nvngcag,1950,0.0168\n\nvngistemp,1950,0.01\n\nvngcag,1949,0.03\n\nvngistemp,1949,0.04\n\nvngcag,1948,0.04\n\nvngcag,1947,0.04\n\nvngcag,1946,0.04\n\nvngcag,1945,0.17\n\nvngcag,1944,0.12\n\nvngcag,1943,0.12\n\nvngcag,1942,0.13\n\nvngcag,1941,0.13\n\nvngcag,1940,0.08\n\nvngcag,1939,0.08\n\nvngcag,1938,0.08\n\nvngcag,1937,0.08\n\nvngcag,1936,0.08\n\nvngcag,1935,0.08\n\nvngcag,1934,0.08\n\nvngcag,1933,0.08\n\nvngcag,1932,0.08\n\nvngcag,1931,0.08\n\nvngcag,1930,0.08\n\nvngcag,1929,0.08\n\nvngcag,1928,0.08\n\nvngcag,1927,0.08\n\nvngcag,1926,0.08\n\nvngcag,1925,0.08\n\nvngcag,1924,0.08\n\nvngcag,1923,0.08\n\nvngcag,1922,0.08\n\nvngcag,1921,0.08\n\nvngcag,1920,0.08\n\nvngcag,1919,0.08\n\nvngcag,1918,0.08\n\nvngcag,1917,0.08\n\nvngcag,1916,0.08\n\nvngcag,1915,0.08\n\nvngcag,1914,0.08\n\nvngcag,1913,0.08\n\nvngcag,1912,0.08\n\nvngcag,1911,0.08\n\nvngcag,1910,0.08\n\nvngcag,1909,0.08\n\nvngcag,1908,0.08\n\nvngcag,1907,0.08\n\nvngcag,1906,0.08\n\nvngcag,1905,0.08\n\nvngcag,1904,0.08\n\nvngcag,1903,0.08\n\nvngcag,1902,0.08\n\nvngcag,1901,0.08\n\nvngcag,1900,0.08\n\nvngcag,1899,0.08\n\nvngcag,1898,0.08\n\nvngcag,1897,0.08\n\nvngcag,1896,0.08\n\nvngcag,1895,0.08\n\nvngcag,1894,0.08\n\nvngcag,1893,0.08\n\nvngcag,1892,0.08\n\nvngcag,1891,0.08\n\nvngcag,1890,0.08\n\nvngcag,1889,0.08\n\nvngcag,1888,0.08\n\nvngcag,1887,0.08\n\nvngcag,1886,0.08\n\nvngcag,1885,0.08\n\nvngcag,1884,0.08\n\nvngcag,1883,0.08\n\nvngcag,1882,0.08\n\nvngcag,1881,0.08\n\nvngcag,1880,0.08\n\nvngcag,1879,0.08\n\nvngcag,1878,0.08\n\nvngcag,1877,0.08\n\nvngcag,1876,0.08\n\nvngcag,1875,0.08\n\nvngcag,1874,0.08\n\nvngcag,1873,0.08\n\nvngcag,1872,0.08\n\nvngcag,1871,0.08\n\nvngcag,1870,0.08\n\nvngcag,1869,0.08\n\nvngcag,1868,0.08\n\nvngcag,1867,0.08\n\nvngcag,1866,0.08\n\nvngcag,1865,0.08\n\nvngcag,1864,0.08\n\nvngcag,1863,0.08\n\nvngcag,1862,0.08\n\nvngcag,1861,0.08\n\nvngcag,1860,0.08\n\nvngcag,1859,0.08\n\nvngcag,1858,0.08\n\nvngcag,1857,0.08\n\nvngcag,1856,0.08\n\nvngcag,1855,0.08\n\nvngcag,1854,0.08\n\nvngcag,1853,0.08\n\nvngcag,1852,0.08\n\nvngcag,1851,0.08\n\nvngcag,1850,0.08\n\nvngcag,1849,0.08\n\nvngcag,1848,0.08\n\nvngcag,1847,0.08\n\nvngcag,1846,0.08\n\nvngcag,1845,0.08\n\nvngcag,1844,0.08\n\nvngcag,1843,0.08\n\nvngcag,1842,0.08\n\nvngcag,1841,0.08\n\nvngcag,1840,0.08\n\nvngcag,1839,0.08\n\nvngcag,1838,0.08\n\nvngcag,1837,0.08\n\nvngcag,1836,0.08\n\nvngcag,1835,0.08\n\nvngcag,1834,0.08\n\nvngcag,1833,0.08\n\nvngcag,1832,0.08\n\nvngcag,1831,0.08\n\nvngcag,1830,0.08\n\nvngcag,1829,0.08\n\nvngcag,1828,0.08\n\nvngcag,1827,0.08\n\nvngcag,1826,0.08\n\nvngcag,1825,0.08\n\nvngcag,1824,0.08\n\nvngcag,1823,0.08\n\nvngcag,1822,0.08\n\nvngcag,1821,0.08\n\nvngcag,1820,0.08\n\nvngcag,1819,0.08\n\nvngcag,1818,0.08\n\nvngcag,1817,0.08\n\nvngcag,1816,0.08\n\nvngcag,1815,0.08\n\nvngcag,1814,0.08\n\nvngcag,1813,0.08\n\nvngcag,1812,0.08\n\nvngcag,1811,0.08\n\nvngcag,1810,0.08\n\nvngcag,1809,0.08\n\nvngcag,1808,0.08\n\nvngcag,1807,0.08\n\nvngcag,1806,0.08\n\nvngcag,1805,0.08\n\nvngcag,1804,0.08\n\nvngcag,1803,0.08\n\nvngcag,1802,0.08\n\nvngcag,1801,0.08\n\nvngcag,1800,0.08\n\nvngcag,1799,0.08\n\nvngcag,1798,0.08\n\nvngcag,1797,0.08\n\nvngcag,1796,0.08\n\nvngcag,1795,0.08\n\nvngcag,1794,0.08\n\nvngcag,1793,0.08\n\nvngcag,1792,0.08\n\nvngcag,1791,0.08\n\nvngcag,1790,0.08\n\nvngcag,1789,0.08\n\nvngcag,1788,0.08\n\nvngcag,1787,0.08\n\nvngcag,1786,0.08\n\nvngcag,1785,0.08\n\nvngcag,1784,0.08\n\nvngcag,1783,0.08\n\nvngcag,1782,0.08\n\nvngcag,1781,0.08\n\nvngcag,1780,0.08\n\nvngcag,1779,0.08\n\nvngcag,1778,0.08\n\nvngcag,1777,0.08\n\nvngcag,1776,0.08\n\nvngcag,1775,0.08\n\nvngcag,1774,0.08\n\nvngcag,1773,0.08\n\nvngcag,1772,0.08\n\nvngcag,1771,0.08\n\nvngcag,1770,0.08\n\nvngcag,1769,0.08\n\nvngcag,1768,0.08\n\nvngcag,1767,0.08\n\nvngcag,1766,0.08\n\nvngcag,1765,0.08\n\nvngcag,1764,0.08\n\nvngcag,1763,0.08\n\nvngcag,1762,0.08\n\nvngcag,1761,0.08\n\nvngcag,1760,0.08\n\nvngcag,1759,0.08\n\nvngcag,1758,0.08\n\nvngcag,1757,0.08\n\nvngcag,1756,0.08\n\nvngcag,1755,0.08\n\nvngcag,1754,0.08\n\nvngcag,1753,0.08\n\nvngcag,1752,0.08\n\nvngcag,1751,0.08\n\nvngcag,1750,0.08\n\nvngcag,1749,0.08\n\nvngcag,1748,0.08\n\nvngcag,1747,0.08\n\nvngcag,1746,0.08\n\nvngcag,1745,0.08\n\nvngcag,1744,0.08\n\nvngcag,1743,0.08\n\nvngcag,1742,0.08\n\nvngcag,1741,0.08\n\nvngcag,1740,0.08\n\nvngcag,1739,0.08\n\nvngcag,1738,0.08\n\nvngcag,1737,0.08\n\nvngcag,1736,0.08\n\nvngcag,1735,0.08\n\nvngcag,1734,0.08\n\nvngcag,1733,0.08\n\nvngcag,1732,0.08\n\nvngcag,1731,0.08\n\nvngcag,1730,0.08\n\nvngcag,1729,0.08\n\nvngcag,1728,0.08\n\nvngcag,1727,0.08\n\nvngcag,1726,0.08\n\nvngcag,1725,0.08\n\nvngcag,1724,0.08\n\nvngcag,1723,0.08\n\nvngcag,1722,0.08\n\nvngcag,1721,0.08\n\nvngcag,1720,0.08\n\nvngcag,1719,0.08\n\nvngcag,1718,0.08\n\nvngcag,1717,0.08\n\nvngcag,1716,0.08\n\nvngcag,1715,0.08\n\nvngcag,1714,0.08\n\nvngcag,1713,0.08\n\nvngcag,1712,0.08\n\nvngcag,1711,0.08\n\nvngcag,1710,0.08\n\nvngcag,1709,0.08\n\nvngcag,1708,0.08\n\nvngcag,1707,0.08\n\nvngcag,1706,0.08\n\nvngcag,1705,0.08\n\nvngcag,1704,0.08\n\nvngcag,1703,0.08\n\nvngcag,1702,0.08\n\nvngcag,1701,0.08\n\nvngcag,1700,0.08\n\nvngcag,1699,0.08\n\nvngcag,1698,0.08\n\nvngcag,1697,0.08\n\nvngcag,1696,0.08\n\nvngcag,1695,0.08\n\nvngcag,1694,0.08\n\nvngcag,1693,0.08\n\nvngcag,1692,0.08\n\nvngcag,1691,0.08\n\nvngcag,1690,0.08\n\nvngcag,1689,0.08\n\nvngcag,1688,0.08\n\nvngcag,1687,0.08\n\nvngcag,1686,0.08\n\nvngcag,1685,0.08\n\nvngcag,1684,0.08\n\nvngcag,1683,0.08\n\nvngcag,1682,0.08\n\nvngcag,1681,0.08\n\nvngcag,1680,0.08\n\nvngcag,1679,0.08\n\nvngcag,1678,0.08\n\nvngcag,1677,0.08\n\nvngcag,1676,0.08\n\nvngcag,1675,0.08\n\nvngcag,1674,0.08\n\nvngcag,1673,0.08\n\nvngcag,1672,0.08\n\nvngcag,1671,0.08\n\nvngcag,1670,0.08\n\nvngcag,1669,0.08\n\nvngcag,1668,0.08\n\nvngcag,1667,0.08\n\nvngcag,1666,0.08\n\nvngcag,1665,0.08\n\nvngcag,1664,0.08\n\nvngcag,1663,0.08\n\nvngcag,1662,0.08\n\nvngcag,1661,0.08\n\nvngcag,1660,0.08\n\nvngcag,1659,0.08\n\nvngcag,1658,0.08\n\nvngcag,1657,0.08\n\nvngcag,1656,0.08\n\nvngcag,1655,0.08\n\nvngcag,1654,0.08\n\nvngcag,1653,0.08\n\nvngcag,1652,0.08\n\nvngcag,1651,0.08\n\nvngcag,1650,0.08\n\nvngcag,1649,0.08\n\nvngcag,1648,0.08\n\nvngcag,1647,0.08\n\nvngcag,1646,0.08\n\nvngcag,1645,0.08\n\nvngcag,1644,0.08\n\nvngcag,1643,0.08\n\nvngcag,1642,0.08\n\nvngcag,1641,0.08\n\nvngcag,1640,0.08\n\nvngcag,1639,0.08\n\nvngcag,1638,0.08\n\nvngcag,1637,0.08\n\nvngcag,1636,0.08\n\nvngcag,1635,0.08\n\nvngcag,1634,0.08\n\nvngcag,1633,0.08\n\nvngcag,1632,0.08\n\nvngcag,1631,0.08\n\nvngcag,1630,0.08\n\nvngcag,1629,0.08\n\nvngcag,1628,0.08\n\nvngcag,1627,0.08\n\nvngcag,1626,0.08\n\nvngcag,1625,0.08\n\nvngcag,1624,0.08\n\nvngcag,1623,0.08\n\nvngcag,1622,0.08\n\nvngcag,1621,0.08\n\nvngcag,1620,0.08\n\nvngcag,1619,0.08\n\nvngcag,1618,0.08\n\nvngcag,1617,0.08\n\nvngcag,1616,0.08\n\nvngcag,1615,0.08\n\nvngcag,1614,0.08\n\nvngcag,1613,0.08\n\nvngcag,1612,0.08\n\nvngcag,1611,0.08\n\nvngcag,1610,0.08\n\nvngcag,1609,0.08\n\nvngcag,1608,0.08\n\nvngcag,1607,0.08\n\nvngcag,1606,0.08\n\nvngcag,1605,0.08\n\nvngcag,1604,0.08\n\nvngcag,1603,0.08\n\nvngcag,1602,0.08\n\nvngcag,1601,0.08\n\nvngcag,1600,0.08\n\nvngcag,1599,0.08\n\nvngcag,1598,0.08\n\nvngcag,1597,0.08\n\nvngcag,1596,0.08\n\nvngcag,1595,0.08\n\nvngcag,1594,0.08\n\nvngcag,1593,0.08\n\nvngcag,1592,0.08\n\nvngcag,1591,0.08\n\nvngcag,1590,0.08\n\nvngcag,1589,0.08\n\nvngcag,1588,0.08\n\nvngcag,1587,0.08\n\nvngcag,1586,0.08\n\nvngcag,1585,0.08\n\nvngcag,1584,0.08\n\nvngcag,1583,0.08\n\nvngcag,1582,0.08\n\nvngcag,1581,0.08\n\nvngcag,1580,0.08\n\nvngcag,1579,0.08\n\nvngcag,1578,0.08\n\nvngcag,1577,0.08\n\nvngcag,1576,0.08\n\nvngcag,1575,0.08\n\nvngcag,1574,0.08\n\nvngcag,1573,0.08\n\nvngcag,1572,0.08\n\nvngcag,1571,0.08\n\nvngcag,1570,0.08\n\nvngcag,1569,0.08\n\nvngcag,1568,0.08\n\nvngcag,1567,0.08\n\nvngcag,1566,0.08\n\nvngcag,1565,0.08\n\nvngcag,1564,0.08\n\nvngcag,1563,0.08\n\nvngcag,1562,0.08\n\nvngcag,1561,0.08\n\nvngcag,1560,0.08\n\nvngcag,1559,0.08\n\nvngcag,1558,0.08\n\nvngcag,1557,0.08\n\nvngcag,1556,0.08\n\nvngcag,1555,0.08\n\nvngcag,1554,0.08\n\nvngcag,1553,0.08\n\nvngcag,1552,0.08\n\nvngcag,1551,0.08\n\nvngcag,1550,0.08\n\nvngcag,1549,0.08\n\nvngcag,1548,0.08\n\nvngcag,1547,0.08\n\nvngcag,1546,0.08\n\nvngcag,1545,0.08\n\nvngcag,1544,0.08\n\nvngcag,1543,0.08\n\nvngcag,1542,0.08\n\nvngcag,1541,0.08\n\nvngcag,1540,0.08\n\nvngcag,1539,0.08\n\nvngcag,1538,0.08\n\nvngcag,1537,0.08\n\nvngcag,1536,0.08\n\nvngcag,1535,0.08\n\nvngcag,1534,0.08\n\nvngcag,1533,0.08\n\nvngcag,1532,0.08\n\nvngcag,1531,0.08\n\nvngcag,1530,0.08\n\nvngcag,1529,0.08\n\nvngcag,1528,0.08\n\nvngcag,1527,0.08\n\nvngcag,1526,0.08\n\nvngcag,1525,0.08\n\nvngcag,1524,0.08\n\nvngcag,1523,0.08\n\nvngcag,1522,0.08\n\nvngcag,1521,0.08\n\nvngcag,1520,0.08\n\nvngcag,1519,0.08\n\nvngcag,1518,0.08\n\nvngcag,1517,0.08\n\nvngcag,1516,0.08\n\nvngcag,1515,0.08\n\nvngcag,1514,0.08\n\nvngcag,1513,0.08\n\nvngcag,1512,0.08\n\nvngcag,1511,0.08\n\nvngcag,1510,0.08\n\nvngcag,1509,0.08\n\nvngcag,1508,0.08\n\nvngcag,1507,0.08\n\nvngcag,1506,0.08\n\nvngcag,1505,0.08\n\nvngcag,1504,0.08\n\nvngcag,1503,0.08\n\nvngcag,1502,0.08\n\nvngcag,1501,0.08\n\nvngcag,1500,0.08\n\nvngcag,1499,0.08\n\nvngcag,1498,0.08\n\nvngcag,1497,0.08\n\nvngcag,1496,0.08\n\nvngcag,1495,0.08\n\nvngcag,1494,0.08\n\nvngcag,1493,0.08\n\nvngcag,1492,0.08\n\nvngcag,1491,0.08\n\nvngcag,1490,0.08\n\nvngcag,1489,0.08\n\nvngcag,1488,0.08\n\nvngcag,1487,0.08\n\nvngcag,1486,0.08\n\nvngcag,1485,0.08\n\nvngcag,1484,0.08\n\nvngcag,1483,0.08\n\nvngcag,1482,0.08\n\nvngcag,1481,0.08\n\nvngcag,1480,0.08\n\nvngcag,1479,0.08\n\nvngcag,1478,0.08\n\nvngcag,1477,0.08\n\nvngcag,1476,0.08\n\nvngcag,1475,0.08\n\nvngcag,1474,0.08\n\nvngcag,1473,0.08\n\nvngcag,1472,0.08\n\nvngcag,1471,0.08\n\nvngcag,1470,0.08\n\nvngcag,1469,0.08\n\nvngcag,1468,0.08\n\nvngcag,1467,0.08\n\nvngcag,1466,0.08\n\nvngcag,1465,0.08\n\nvngcag,1464,0.08\n\nvngcag,1463,0.08\n\nvngcag,1462,0.08\n\nvngcag,1461,0.08\n\nvngcag,1460,0.08\n\nvngcag,1459,0.08\n\nvngcag,1458,0.08\n\nvngcag,1457,0.08\n\nvngcag,1456,0.08\n\nvngcag,1455,0.08\n\nvngcag,1454,0.08\n\nvngcag,1453,0.08\n\nvngcag,1452,0.08\n\nvngcag,1451,0.08\n\nvngcag,1450,0.08\n\nvngcag,1449,0.08\n\nvngcag,1448,0.08\n\nvngcag,1447,0.08\n\nvngcag,1446,0.08\n\nvngcag,1445,0.08\n\nvngcag,1444,0.08\n\nvngcag,1443,0.08\n\nvngcag,1442,0.08\n\nvngcag,1441,0.08\n\nvngcag,1440,0.08\n\nvngcag,1439,0.08\n\nvngcag,1438,0.08\n\nvngcag,1437,0.08\n\nvngcag,1436,0.08\n\nvngcag,1435,0.08\n\nvngcag,1434,0.08\n\nvngcag,1433,0.08\n\nvngcag,1432,0.08\n\nvngcag,1431,0.08\n\nvngcag,1430,0.08\n\nvngcag,1429,0.08\n\nvngcag,1428,0.08\n\nvngcag,1427,0.08\n\nvngcag,1426,0.08\n\nvngcag,1425,0.08\n\nvngcag,1424,0.08\n\nvngcag,1423,0.08\n\nvngcag,1422,0.08\n\nvngcag,1421,0.08\n\nvngcag,1420,0.08\n\nvngcag,1419,0.08\n\nvngcag,1418,0.08\n\nvngcag,1417,0.08\n\nvngcag,1416,0.08\n\nvngcag,1415,0.08\n\nvngcag,1414,0.08\n\nvngcag,1413,0.08\n\nvngcag,1412,0.08\n\nvngcag,1411,0.08\n\nvngcag,1410,0.08\n\nvngcag,1409,0.08\n\nvngcag,1408,0.08\n\nvngcag,1407,0.08\n\nvngcag,1406,0.08\n\nvngcag,1405,0.08\n\nvngcag,1404,0.08\n\nvngcag,1403,0.08\n\nvngcag,1402,0.08\n\nvngcag,1401,0.08\n\nvngcag,1400,0.08\n\nvngcag,1399,0.08\n\nvngcag,1398,0.08\n\nvngcag,1397,0.08\n\nvngcag,1396,0.08\n\nvngcag,1395,0.08\n\nvngcag,1394,0.08\n\nvngcag,1393,0.08\n\nvngcag,1392,0.08\n\nvngcag,1391,0.08\n\nvngcag,1390,0.08\n\nvngcag,1389,0.08\n\nvngcag,1388,0.08\n\nvngcag,1387,0.08\n\nvngcag,1386,0.08\n\nvngcag,1385,0.08\n\nvngcag,1384,0.08\n\nvngcag,1383,0.08\n\nvngcag,1382,0.08\n\nvngcag,1381,0.08\n\nvngcag,1380,0.08\n\nvngcag,1379,0.08\n\nvngcag,1378,0.08\n\nvngcag,1377,0.08\n\nvngcag,1376,0.08\n\nvngcag,1375,0.08\n\nvngcag,1374,0.08\n\nvngcag,1373,0.08\n\nvngcag,1372,0.08\n
```


1880	-0.1148	0.823953	False
1881	-0.0628	0.180163	False
1882	-0.0648	0.13378	False
1883	-0.1424	0.146379	False
1884	-0.2009	0.882296	False

```
In [ ]: # just data from one year from just the "Mean" column
df.loc[1880:1884, 'Mean']
```

```
Out[ ]: 0.9363
```

```
In [ ]: # a range of years
df.loc[1880:1884, 'Mean']
```

```
Out[ ]: Year
1880    -0.1148
1881    -0.0628
1882    -0.0648
1883    -0.1424
1884    -0.2009
Name: Mean, dtype: float64
```

```
In [ ]: # range of rows and columns
df.loc[1880:1884, ['Mean', 'HighLow']]
```

```
Out[ ]:      Mean  HighLow
Year
1880    -0.1148    False
1881    -0.0628    False
1882    -0.0648    False
1883    -0.1424    False
1884    -0.2009    False
```

```
In [ ]: # note that the order in which you ask for columns impacts the output
row_ind = [1880, 1980]
col_ind = ['HighLow', 'Mean']
df2 = df.loc[row_ind, col_ind]
df2.head()
```

```
Out[ ]:      HighLow  Mean
Year
1880      False  -0.1148
1980       True   0.2637
```

A few more examples of using iloc to index into specific rows, columns

```
In [54]: df.head()
```

```
Out[54]:      Mean      Std
Year
2016  0.9363  0.751983
2015      NaN      NaN
2014  0.7408  0.308372
2013  0.6679  0.064226
2012  0.6240  0.369924
```

```
In [58]: df.iloc[:20:2]
```

```
Out[58]:      Mean      Std
Year
2016  0.9363  0.751983
2014  0.7408  0.308372
2012  0.6240  0.369924
2010  0.7014  0.006540
2008  0.5419  0.135968
2006  0.6125  0.477445
2004  0.5783  0.350637
2002  0.6023  0.958127
2000  0.4262  0.315162
1998  0.6344  0.318651
```

```
In [63]: df.iloc[:10, 0]
```

```
Out[63]: Year
2016    0.9363
2015      NaN
2014    0.7408
2013    0.6679
2012    0.6240
2011    0.5788
2010    0.7014
2009    0.6367
2008    0.5419
2007    0.6100
Name: Mean, dtype: float64
```

Deleting columns...

```
In [68]: df_temp = df.copy()
```

```
In [69]: df_temp.head()
```

```
Out[69]:      Mean      Std
Year
2016  0.9363  0.751983
2015      NaN      NaN
2014  0.7408  0.308372
2013  0.6679  0.064226
2012  0.6240  0.369924
```

```
In [70]: # using the del command will delete a column from the DF
# make an explicit copy! (why???) so we don't overwrite our original data frame
df2 = df_temp.copy()
del df2['Mean']
df2.head()
```

```
Out[70]:      Std
Year
2016  0.751983
2015      NaN
2014  0.308372
2013  0.064226
2012  0.369924
```

```
In [71]: df_temp.head()
```

```
Out[71]:      Mean      Std
Year
2016  0.9363  0.751983
2015      NaN      NaN
2014  0.7408  0.308372
2013  0.6679  0.064226
2012  0.6240  0.369924
```

Replace method

- first param is what you want to replace (can be any data type as needed)
- second param is what you want to replace it with

```
In [72]: # make a df from a dictionary
data = {'D1': [6,8,4,3], 'D2': [4,4,2,1]}
```

```
df = pd.DataFrame(data)
df.head()
```

```
Out[72]:      D1  D2
0     6   4
1     8   4
2     4   2
3     3   1
```

```
In [75]: df = df.replace(4, 7)
df.head()
```

```
Out[75]:      D1  D2
0     6   7
1     8   7
2     7   2
3     3   1
```

Set index from a column in a data frame

```
In [79]: data = {'D1': [6,8,4,3], 'D2': [4,4,2,1], 'D3': [2018,2019,2020,2021]}
```

```
df = pd.DataFrame(data)
df.head()
```

```
Out[79]:      D1  D2  D3
2018     6   4  2018
2019     8   4  2019
2020     4   2  2020
2021     3   1  2021
```

```
In [78]: df = df.set_index('D3')
df.head()
```

```
Out[78]:      D1  D2
D3
2018     6   4
2019     8   4
2020     4   2
2021     3   1
```

Loop over data frame and compute the mean of 'w' consecutive rows

```
In [86]: # make a df
import random as random
random.seed(10)
n_data_pts = 100
d1=[]
d2=[]
for i in range(n_data_pts):
    d1.append(random.randint(0,40))
    d2.append(random.randint(0,30))
# then make a data frame
df = pd.DataFrame({'D1':d1, 'D2':d2})
df.head()
```

```
Out[86]:      D1  D2
0    36   1
1    27  15
2    36   0
3    13  26
```

```
In [100]: win = 20 # moving average window
n = len(df)
# init a list to append moving average
m_avg = []
# init a counter to keep track of where we are in the DF
for i in range(0,n-win):
    m_avg.append(df.loc[i:i+win, 'D1'].mean())
# print out our list of windowed averages
print(len(m_avg))
print(m_avg)
[21.714285714285715, 17.952380952380953, 17.476190476190474, 22.61904761904762, 21.1]
5
```

Preview of next week - quick intro to plotting

```
In [94]: import matplotlib.pyplot as plt
```

```
In [101]: plt.plot(df['D1'])
plt.show()
```

```
plt.plot(m_avg)
plt.show()
```



```
In [98]:
```