Name:Viviyan Richards W

Roll no:205229133

In [1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

In [2]:
```
north_america = pd.read_csv('north_america_2000_2010.csv', index_col=0)
```

In [3]:
```
south_america = pd.read_csv('south_america_2000_2010.csv', index_col=0)
```

In [4]:
```
north_america
```

Out[4]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Canada	1779.0	1771.0	1754.0	1740.0	1760.0	1747	1745.0	1741.0	1735	1701.0	1703.0
Mexico	2311.2	2285.2	2271.2	2276.5	2270.6	2281	2280.6	2261.4	2258	2250.2	2242.4
USA	1836.0	1814.0	1810.0	1800.0	1802.0	1799	1800.0	1798.0	1792	1767.0	1778.0

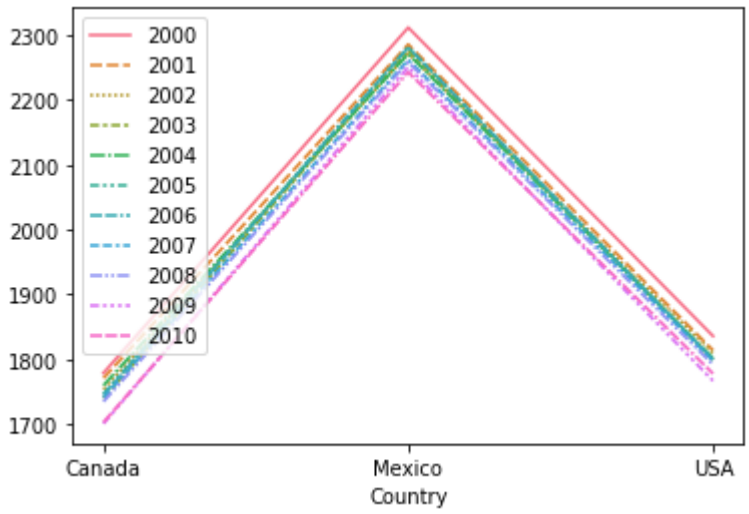
In [5]:
```
south_america
```

Out[5]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Chile	2263	2242	2250	2235	2232	2157	2165	2128	2095	2074	2069.6

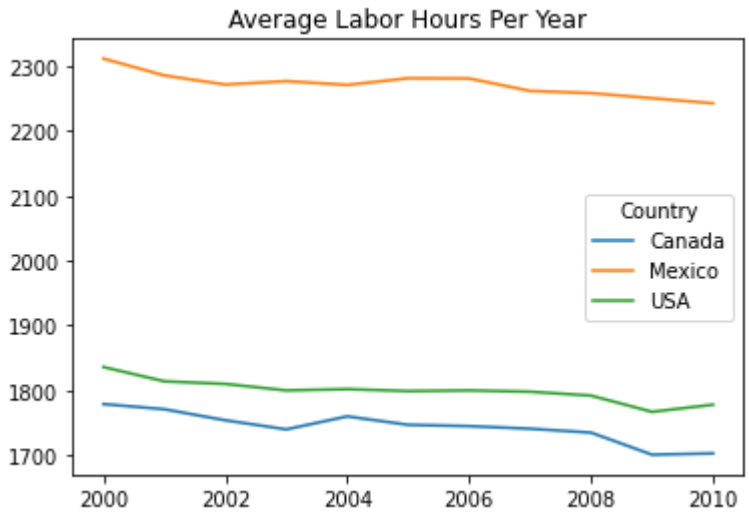
In [6]:
```
sns.lineplot(data=north_america)
```

Out[6]:
```
<AxesSubplot:xlabel='Country'>
```



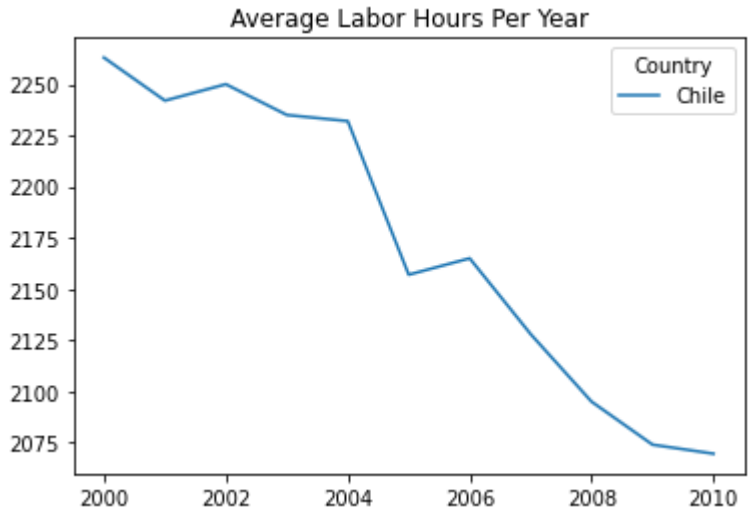
In [7]:
```
north_america.T.plot(title="Average Labor Hours Per Year")
```

Out[7]:
```
<AxesSubplot:title={'center':'Average Labor Hours Per Year'}>
```



In [8]:
```
south_america.T.plot(title="Average Labor Hours Per Year")
```

Out[8]:
```
<AxesSubplot:title={'center':'Average Labor Hours Per Year'}>
```



In [9]:
```
americas = pd.concat([north_america,south_america])
```

In [10]:
```
americas
```

Out[10]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Canada	1779.0	1771.0	1754.0	1740.0	1760.0	1747	1745.0	1741.0	1735	1701.0	1703.0
Mexico	2311.2	2285.2	2271.2	2276.5	2270.6	2281	2280.6	2261.4	2258	2250.2	2242.4
USA	1836.0	1814.0	1810.0	1800.0	1802.0	1799	1800.0	1798.0	1792	1767.0	1778.0
Chile	2263.0	2242.0	2250.0	2235.0	2232.0	2157	2165.0	2128.0	2095	2074.0	2069.6

In [11]:
```
america_1 = pd.read_csv('americas_2011.csv', index_col=0)
america_2 = pd.read_csv('americas_2012.csv', index_col=0)
america_3 = pd.read_csv('americas_2013.csv', index_col=0)
america_4 = pd.read_csv('americas_2014.csv', index_col=0)
america_5 = pd.read_csv('americas_2015.csv', index_col=0)
```

In [12]:
```
temp=america_1.join(america_2)
```

In [13]:
```
temp=temp.join(america_3)
```

In [14]:
```
temp=temp.join(america_4)
```

In [15]:
```
temp=temp.join(america_5)
```

In [16]:
```
temp
```

Out[16]:

Country	2011	2012	2013	2014	2015
Canada	1700.0	1713.0	1707.0	1703.0	1706.0
Chile	2047.4	2024.0	2015.3	1990.1	1987.5
Mexico	2250.2	2225.8	2236.6	2228.4	2246.4
USA	1786.0	1789.0	1787.0	1789.0	1790.0

In [17]:
```
americas = americas.join(temp)
```

In [18]:
```
americas.index.names = ['Country']
```

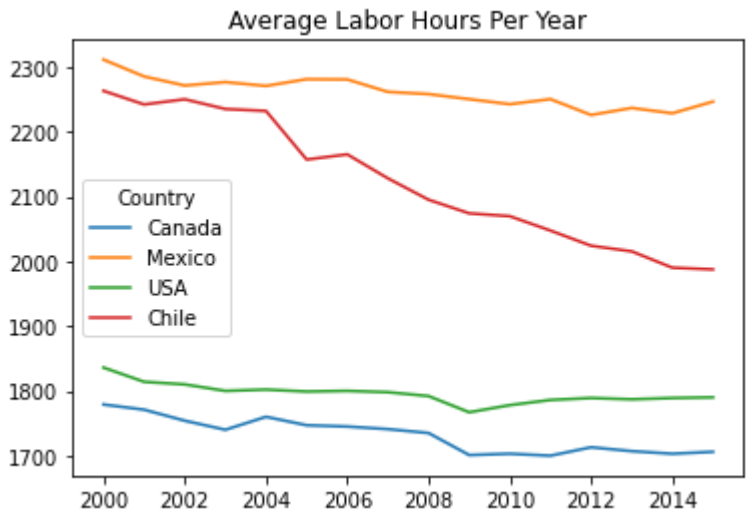
In [19]:
```
americas
```

Out[19]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Canada	1779.0	1771.0	1754.0	1740.0	1760.0	1747	1745.0	1741.0	1735	1701.0	1703.0	1700.0
Mexico	2311.2	2285.2	2271.2	2276.5	2270.6	2281	2280.6	2261.4	2258	2250.2	2242.4	2250.2
USA	1836.0	1814.0	1810.0	1800.0	1802.0	1799	1800.0	1798.0	1792	1767.0	1778.0	1786.0
Chile	2263.0	2242.0	2250.0	2235.0	2232.0	2157	2165.0	2128.0	2095	2074.0	2069.6	2047.4

In [20]:
```
americas.T.plot(title="Average Labor Hours Per Year")
```

Out[20]:
```
<AxesSubplot:title={'center':'Average Labor Hours Per Year'}>
```



In [21]:
```
asia = pd.read_csv('asia_2000_2015.csv', index_col=0)
asia
```

Out[21]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	20
Israel	2017	1979	1993	1974	1942	1931	1919	1931	1929	1927	1918	1920	1910	1867	18
Japan	1821	1809	1798	1799	1787	1775	1784	1785	1771	1714	1733	1728	1745	1734	17
Korea	2512	2499	2464	2424	2392	2351	2346	2306	2246	2232	2187	2090	2163	2079	2'
Russia	1982	1980	1982	1993	1993	1989	1998	1999	1997	1974	1976	1979	1982	1980	19

In [22]:
```
europe = pd.read_csv('europe_2000_2015.csv', index_col=0)
europe.head()
```

Out[22]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Austria	1807.4	1794.6	1792.2	1783.8	1786.8	1764.0	1746.2	1736.0	1728.5	1673.0	1668.6
Belgium	1595.0	1588.0	1583.0	1578.0	1573.0	1565.0	1572.0	1577.0	1570.0	1548.0	1546.0
Switzerland	1673.6	1635.0	1614.0	1626.8	1656.5	1651.7	1643.2	1632.7	1623.1	1614.9	1612.4
Czech Republic	1896.0	1818.0	1816.0	1806.0	1817.0	1817.0	1799.0	1784.0	1790.0	1779.0	1800.0
Germany	1452.0	1441.9	1430.9	1424.8	1422.2	1411.3	1424.7	1424.4	1418.4	1372.7	1389.9

In [23]:
```
south_pacific = pd.read_csv('south_pacific_2000_2015.csv', index_col=0)
south_pacific
```

Out[23]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	20
Australia	1778.7	1736.7	1731.7	1735.8	1734.5	1729.2	1720.5	1712.5	1717.2	1690	1691.5	169
New Zealand	1836.0	1825.0	1826.0	1823.0	1830.0	1815.0	1795.0	1774.0	1761.0	1740	1755.0	174

In [24]:
```
world = pd.concat([americas,europe,south_pacific,asia])
world
```

Out[24]:

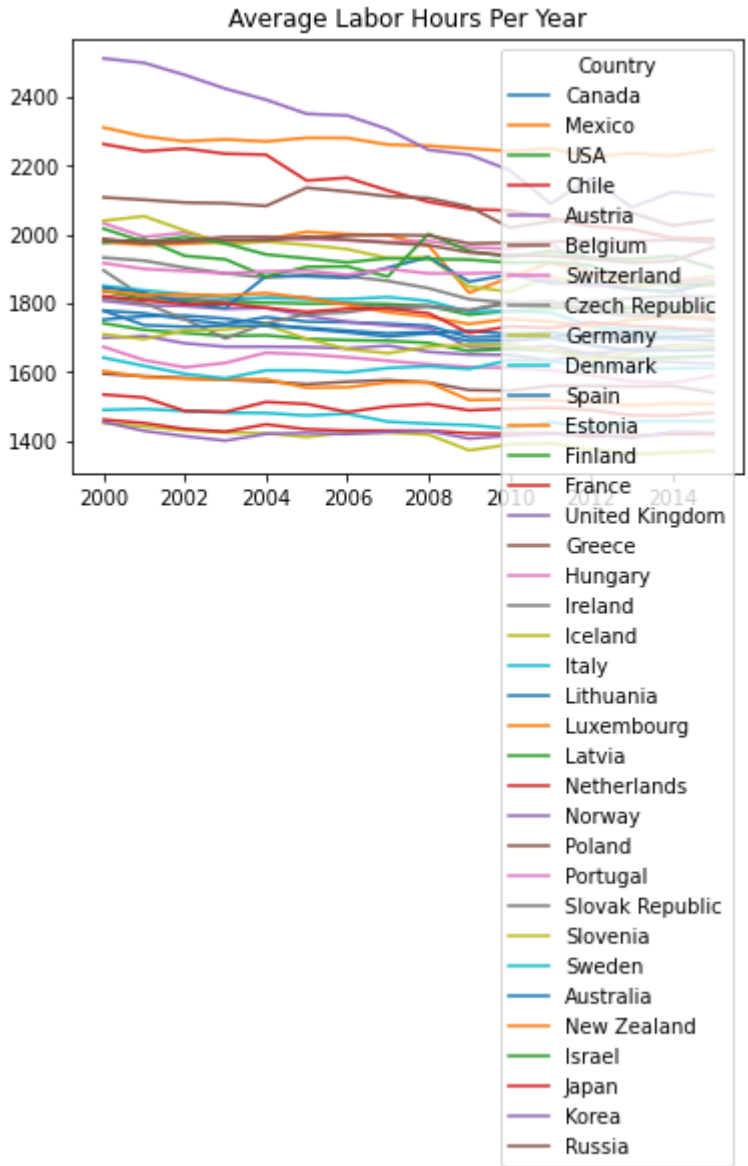
Country	2000	2001	2002	2003	2004	2005	2
Canada	1779.000000	1771.00000	1754.000000	1740.000000	1760.000000	1747.000000	174
Mexico	2311.200000	2285.20000	2271.200000	2276.500000	2270.600000	2281.000000	228
USA	1836.000000	1814.00000	1810.000000	1800.000000	1802.000000	1799.000000	180
Chile	2263.000000	2242.00000	2250.000000	2235.000000	2232.000000	2157.000000	216
Austria	1807.400000	1794.60000	1792.200000	1783.800000	1786.800000	1764.000000	174
Belgium	1595.000000	1588.00000	1583.000000	1578.000000	1573.000000	1565.000000	157
Switzerland	1673.600000	1635.00000	1614.000000	1626.800000	1656.500000	1651.700000	164
Czech Republic	1896.000000	1818.00000	1816.000000	1806.000000	1817.000000	1817.000000	179
Germany	1452.000000	1441.90000	1430.900000	1424.800000	1422.200000	1411.300000	142

In [25]:
```
world.index
```

Out[25]:
```
Index(['Canada', 'Mexico', 'USA', 'Chile', 'Austria', 'Belgium', 'Switzerland',
       'Czech Republic', 'Germany', 'Denmark', 'Spain', 'Estonia', 'Finland',
       'France', 'United Kingdom', 'Greece', 'Hungary', 'Ireland', 'Iceland',
       'Italy', 'Lithuania', 'Luxembourg', 'Latvia', 'Netherlands', 'Norway',
       'Poland', 'Portugal', 'Slovak Republic', 'Slovenia', 'Sweden',
       'Australia', 'New Zealand', 'Israel', 'Japan', 'Korea', 'Russia'],
      dtype='object', name='Country')
```

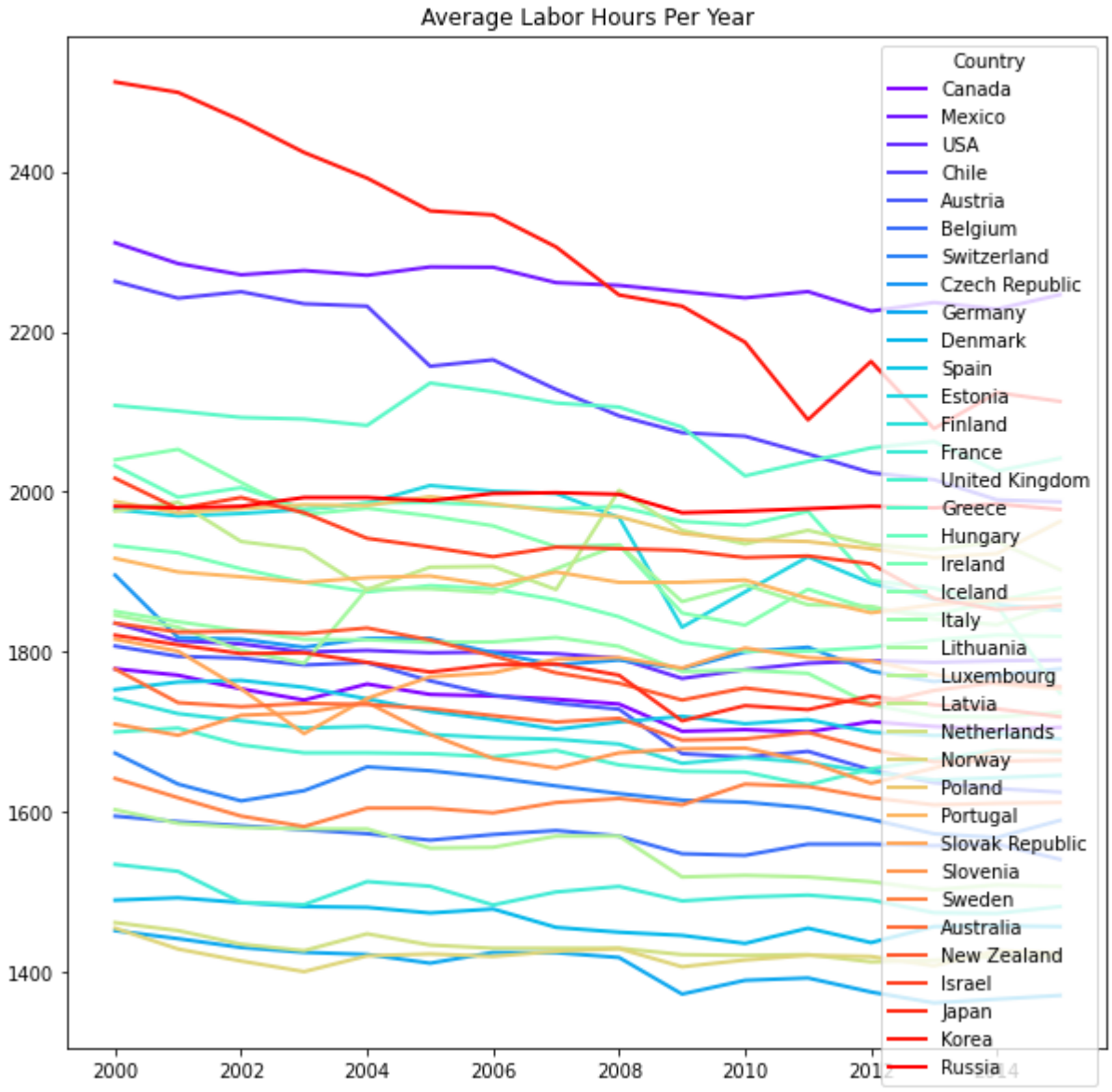
In [27]:
```
world.T.plot(title="Average Labor Hours Per Year")
```

Out[27]:
```
<AxesSubplot:title={'center':'Average Labor Hours Per Year'}>
```



In [28]:
```
world.T.plot(title="Average Labor Hours Per Year",figsize=(10,10),colormap='rainb
```

Out[28]:
```
<AxesSubplot:title={'center':'Average Labor Hours Per Year'}>
```



In [29]:
```
historical = pd.read_csv('historical.csv', index_col=0)
historical.head()
```

Out[29]:

Country	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	...	1990	1991	
Australia	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	1779.5	1774.90	177
Austria	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
Belgium	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	1662.9	1625.79	160
Canada	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	1789.5	1767.50	176
Switzerland	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	1673.10	168

5 rows × 50 columns

In [30]:
```
print("World rows & columns: ", world.shape)
print("Historical rows & columns: ", historical.shape)
```

```
World rows & columns:  (36, 16)
Historical rows & columns:  (39, 50)
```

In [31]:
```
world_historical=world.merge(historical,right_on='Country',left_on='Country')
```

In [32]:
```
world_historical.shape
```

Out[32]:
```
(36, 66)
```

In [33]:
```
world_historical.head()
```

Out[33]:

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	...	1990
Canada	1779.0	1771.0	1754.0	1740.0	1760.0	1747.0	1745.0	1741.0	1735.0	1701.0	...	1789.5
Mexico	2311.2	2285.2	2271.2	2276.5	2270.6	2281.0	2280.6	2261.4	2258.0	2250.2	...	NaN
USA	1836.0	1814.0	1810.0	1800.0	1802.0	1799.0	1800.0	1798.0	1792.0	1767.0	...	1832.0
Chile	2263.0	2242.0	2250.0	2235.0	2232.0	2157.0	2165.0	2128.0	2095.0	2074.0	...	NaN
Austria	1807.4	1794.6	1792.2	1783.8	1786.8	1764.0	1746.2	1736.0	1728.5	1673.0	...	NaN

5 rows × 66 columns

In [34]:
```
world_historial = world.join(historical)
```

In [35]:
```
world_historical.head()
```

Out[35]:

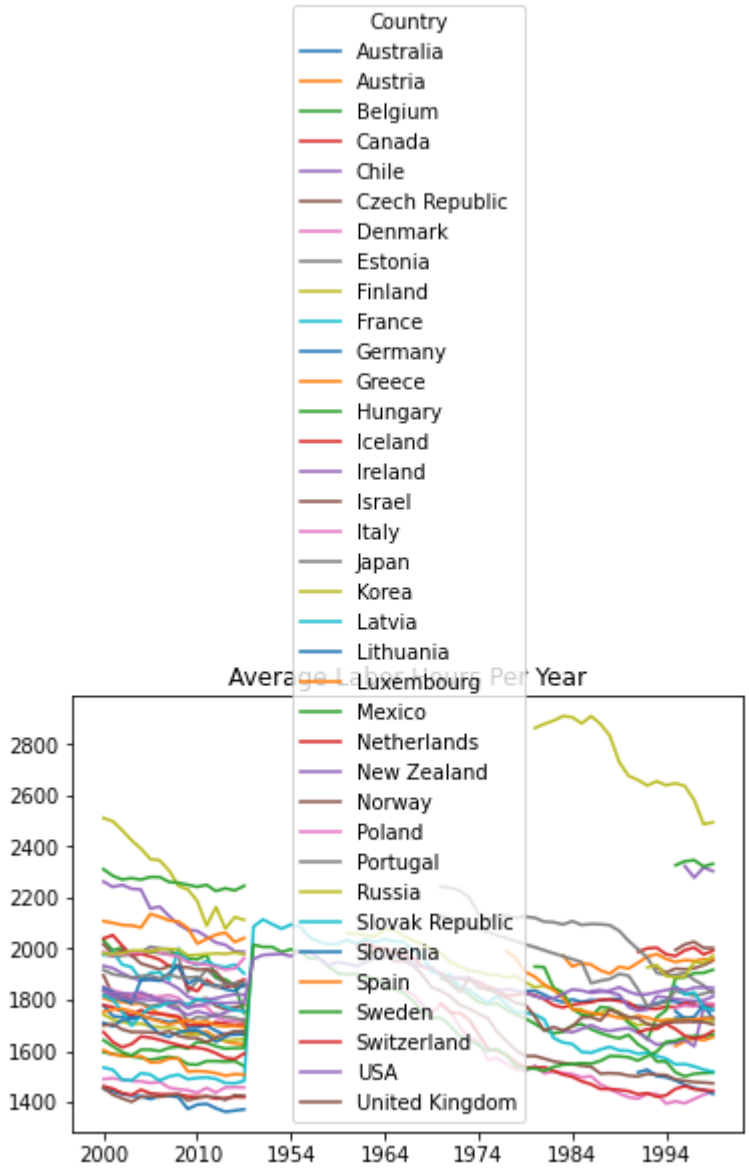
Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	...	1990
Canada	1779.0	1771.0	1754.0	1740.0	1760.0	1747.0	1745.0	1741.0	1735.0	1701.0	...	1789.5
Mexico	2311.2	2285.2	2271.2	2276.5	2270.6	2281.0	2280.6	2261.4	2258.0	2250.2	...	NaN
USA	1836.0	1814.0	1810.0	1800.0	1802.0	1799.0	1800.0	1798.0	1792.0	1767.0	...	1832.0
Chile	2263.0	2242.0	2250.0	2235.0	2232.0	2157.0	2165.0	2128.0	2095.0	2074.0	...	NaN
Austria	1807.4	1794.6	1792.2	1783.8	1786.8	1764.0	1746.2	1736.0	1728.5	1673.0	...	NaN

5 rows × 66 columns

In [36]:
```
world_historical.sort_index(inplace=True)
```

In [37]:
```
world_historical.T.plot(title="Average Labor Hours Per Year")
```

Out[37]: <AxesSubplot:title={'center':'Average Labor Hours Per Year'}>



In [38]:
```
work=world.mean(axis=1)

long=max(world.mean(axis=1))

short=min(world.mean(axis=1))
```

In [39]:
```
print("country worked longer hours per year : ",work[work == long].index[0])
```

```
country worked longer hours per year :  Korea
```

In [40]:
```
print("country worked longer hours per year : ",work[work==short].index[0])
```

```
country worked longer hours per year :  Germany
```