BDA Lab1 - spark

Student: Xuan Wang (xuawa284)

Student: Lepeng Zhang (lepzh903)

**Question1:**

1. Code: lab1-1.py
2. Partial results:

Max:

('1975', 36.1)

('1992', 35.4)

('1994', 34.7)

Min:

('1990', -35.0)

('1952', -35.5)

('1974', -35.6)

('1954', -36.0)

**Question2:**

1. Code: lab1-2.py
2. Partial results:

total\_count:

('1995', '10', 33102)

('2000', '08', 109201)

('2003', '10', 8830)

('2007', '10', 17792)

('2008', '04', 16989)

('2011', '09', 99335)

('1964', '10', 7500)

('1965', '08', 49928)

('1974', '04', 10120)

('1975', '08', 62565)

('1975', '11', 441)

('1976', '08', 60898)

....

distinct\_count:

('1979', '09', 351)

('1982', '10', 315)

('1984', '04', 313)

('1987', '08', 320)

('1989', '08', 316)

('1989', '11', 128)

('1995', '10', 302)

('1996', '04', 263)

('2000', '08', 325)

('2008', '04', 296)

('2011', '09', 315)

('2003', '10', 203)

('1974', '04', 344)

.....

**Question3:**

1. Code: lab1-3.py
2. Partial results:

('1965', '02', '102190', -4.214285714285714)

('1968', '01', '102190', -11.216129032258065)

('1968', '08', '102190', 14.588709677419356)

('1968', '11', '102190', -5.556666666666666)

('1975', '02', '102190', -5.567857142857143)

('1976', '02', '102190', -4.086206896551724)

('1979', '03', '102190', -1.4451612903225806)

('1979', '05', '102190', 7.914516129032258)

('1979', '07', '102190', 13.624193548387096)

('2001', '04', '102190', 2.494999999999999)

....

**Question4:**

1. Code: lab1-4.py
2. Partial results:

Empty, no stations fulfill such requirements.

**Question5:**

1. Code: lab1-5.py
2. Partial results:

('1996', '12', 39.55000000000003)

('1997', '05', 60.80000000000001)

('1998', '10', 53.41666666666671)

('1999', '10', 18.549999999999997)

('2001', '01', 36.416666666666686)

('2003', '01', 17.716666666666672)

('2006', '01', 17.683333333333334)

('2008', '01', 44.9666666666667)

('2008', '03', 42.200000000000024)

('2013', '12', 42.26250000000002)

('2015', '07', 119.09999999999997)

('2015', '11', 63.88750000000002)

('1995', '12', 5.116666666666666)

('1996', '10', 22.45)

...