

BSAD 8700 - Business Analytics

Welcome to Business Analytics!

Instructor: Dusty White, PhD

Who am I?

- Economics Faculty Member
 - Undergrad in Economics, from Brigham Young University
 - PhD in Economics, from Washington State University
 - Data and Data Analytics are my passion!

Contacting me

Email: drwhite@unomaha.edu

Slack: datascienceuno.slack.com

- Use the `#businessanalytics` channel for this class
- Be sure to sign up with your UNO email

I am also happy to arrange meetings by Zoom if you would like a more personal interaction.
We can and will find a time that works for both of us!

Course material

The entire course is accessible through [Canvas!](#)

- You will find lectures here
- You will submit assignments and projects here
- You will find links to all documents here

Software needed for class

- [Tableau Desktop](#) (we will provide a license for the duration of the course!)
- [Orange Data Mining](#)
- [SQLectron](#), a great and free SQL interface
- Microsoft Excel

What do you hope to get out of this course?

Let me know in the Slack channel!

How will this course help you?

- Learn the principles of analytics
- Learn to communicate with Data Analysts
- Learn when Data Analytics should be used
- Understand when Analytics is not being correctly applied.

I'm a Manager, not a Data Scientist...

"It is important to understand data science even if you never intend to apply it yourself. Data-analytic thinking enables you to evaluate proposals for data mining projects. For example, if an employee, a consultant, or a potential investment target proposes to improve a particular business application by extracting knowledge from data, you should be able to assess the proposal systematically and decide whether it is sound or flawed. ... [Y]ou should be able to spot obvious flaws, unrealistic assumptions, and missing pieces."

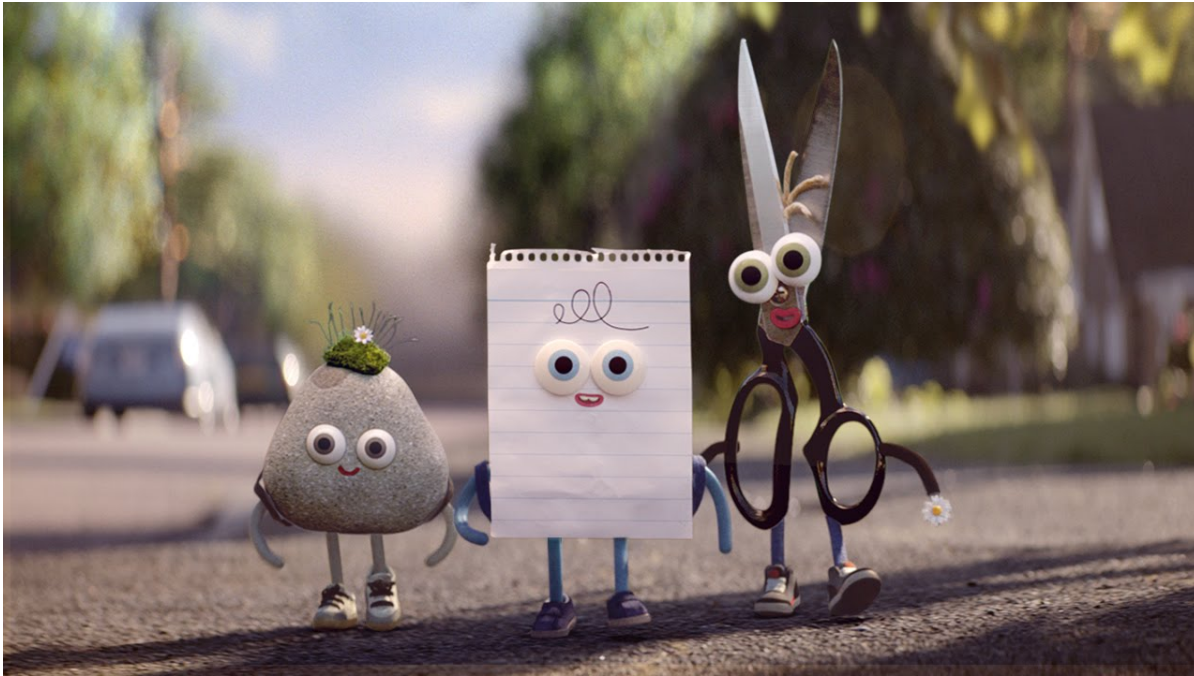
-- Data Science for Business

How can you be successful in this class?

- Do the required readings, and then be active in discussions on Slack!
- Be ready to “get your hands dirty” in lab
- Collaborate with classmates (it's required!)
- Always ask questions!
 - I only know what is interesting to ME
 - Tell me what is interesting to YOU, and I can talk about it

Let's Get Started!

Play a Game



Start by playing a simple game (I recommend about 20 rounds) of [Rock, Paper, Scissors](#)

Where do we actually use models like this?

[The Economist](#) gives us a few examples:

- Predict Repeat Offenders
- Predict Heart Attacks in Hospital Patients
- Predict at-risk Students, Police Officers
- Find children at risk of lead poisoning

Why are these models so valuable to us?

Computers have the ability to "see" lots of data points at once, so if we use algorithms carefully, we can utilize vast amounts of data to inform our decisions.

What is data?

Data is a representation of objects or information from the real world

What is data?

Data is a representation of objects or information from the real world

- Census responses are data

State Illinois County Carle Precinct 1

DEPARTMENT OF COMMERCE, BUREAU OF THE CENSUS
FIFTEENTH CENSUS OF THE UNITED STATES: 1900
POPULATION SCHEDULE

Enumeration District No. 12 of 156
Superior's District No. 2

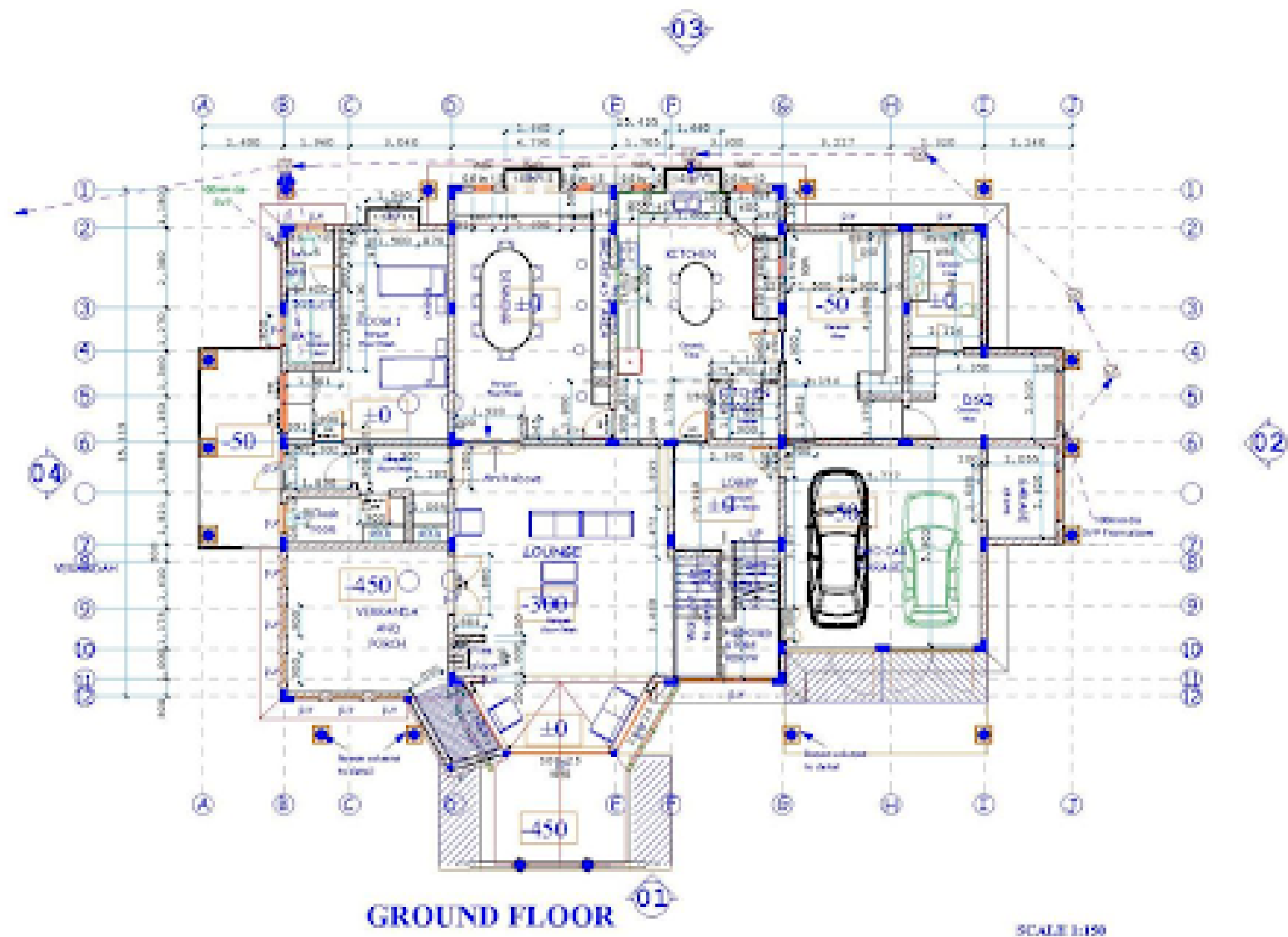
Sheet No. 18

HOUSEHOLD IN DETAIL										PERSONS IN DETAIL									
NAME, SEX, AGE, COLOR, RELATION TO HEAD, MARRIAGE, BIRTHPLACE, NATIVITY, CITIZENSHIP										NAME, SEX, AGE, COLOR, RELATION TO HEAD, MARRIAGE, BIRTHPLACE, NATIVITY, CITIZENSHIP									
NAME, SEX, AGE, COLOR, RELATION TO HEAD, MARRIAGE, BIRTHPLACE, NATIVITY, CITIZENSHIP										NAME, SEX, AGE, COLOR, RELATION TO HEAD, MARRIAGE, BIRTHPLACE, NATIVITY, CITIZENSHIP									
1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019
1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039
1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059
1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079
1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099
1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119
1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139
1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159
1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179
1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199
1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219
1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239
1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259
1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279
1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299
1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319
1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339
1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359
1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379
1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399
1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419
1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439
1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459
1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479
1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499
1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519
1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539
1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559
1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579
1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599
1600	1601	1602	1603	1604	1605	1606	1607	1608	1609	1610	1611	1612	1613	1614	1615	1616	1617	1618	1619
1620	1621	1622	1623	1624	1625	1626	1627	1628	1629	1630	1631	1632	1633	1634	1635	1636	1637	1638	1639
1640	1641	1642	1643	1644	1645	1646	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659
1660	1661	1662	1663	1664	1665	1666	1667	1668	1669	1670	1671	1672	1673	1674	1675	1676	1677	1678	1679
1680	1681	1682	1683	1684	1685	1686	1687	1688	1689	1690	1691	1692	1693	1694	1695	1696	1697	1698	1699
1700	1701	1702	1703	1704	1705	1706	1707	1708	1709	1710	1711	1712	1713	1714	1715	1716	1717	1718	1719
1720	1721	1722	1723	1724	1725	1726	1727	1728	1729	1730	1731	1732	1733	1734	1735	1736	1737	1738	1739
1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759
1760	1761	1762	1763	1764	1765	1766	1767	1768	1769	1770	1771	1772	1773	1774	1775	1776	1777	1778	1779
1780	1781	1782	1783	1784	1785	1786	1787	1788	1789	1790	1791	1792	1793	1794	1795	1796	1797	1798	1799
1800	1801	1802	1803	1804	1805	1806	1807	1808	1809	1810	1811	1812	1813	1814	1815	1816	1817	1818	1819
1820	1821	1822	1823	1824	1825	1826	1827	1828	1829	1830	1831	1832	1833	1834	1835	1836	1837	1838	1839
1840	1841	1842	1843	1844	1845	1846	1847	1848	1849	1850	1851	1852	1853	1854	1855	1856	1857	1858	1859
1860	1861	1862	1863	1864	1865	1866	1867	1868	1869	1870	1871	1872	1873	1874	1875	1876	1877	1878	1879
1880	1881	1882	1883	1884	1885	1886	1887	1888	1889	1890	1891	1892	1893	1894	1895	1896	1897	1898	1899
1900	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919
1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959
1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039
2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059
2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079
2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099
2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119
2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139
2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159
2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179
2180	2181	2182	2183	2184	2185	2186	21												

What is data?

Data is a representation of objects or information from the real world

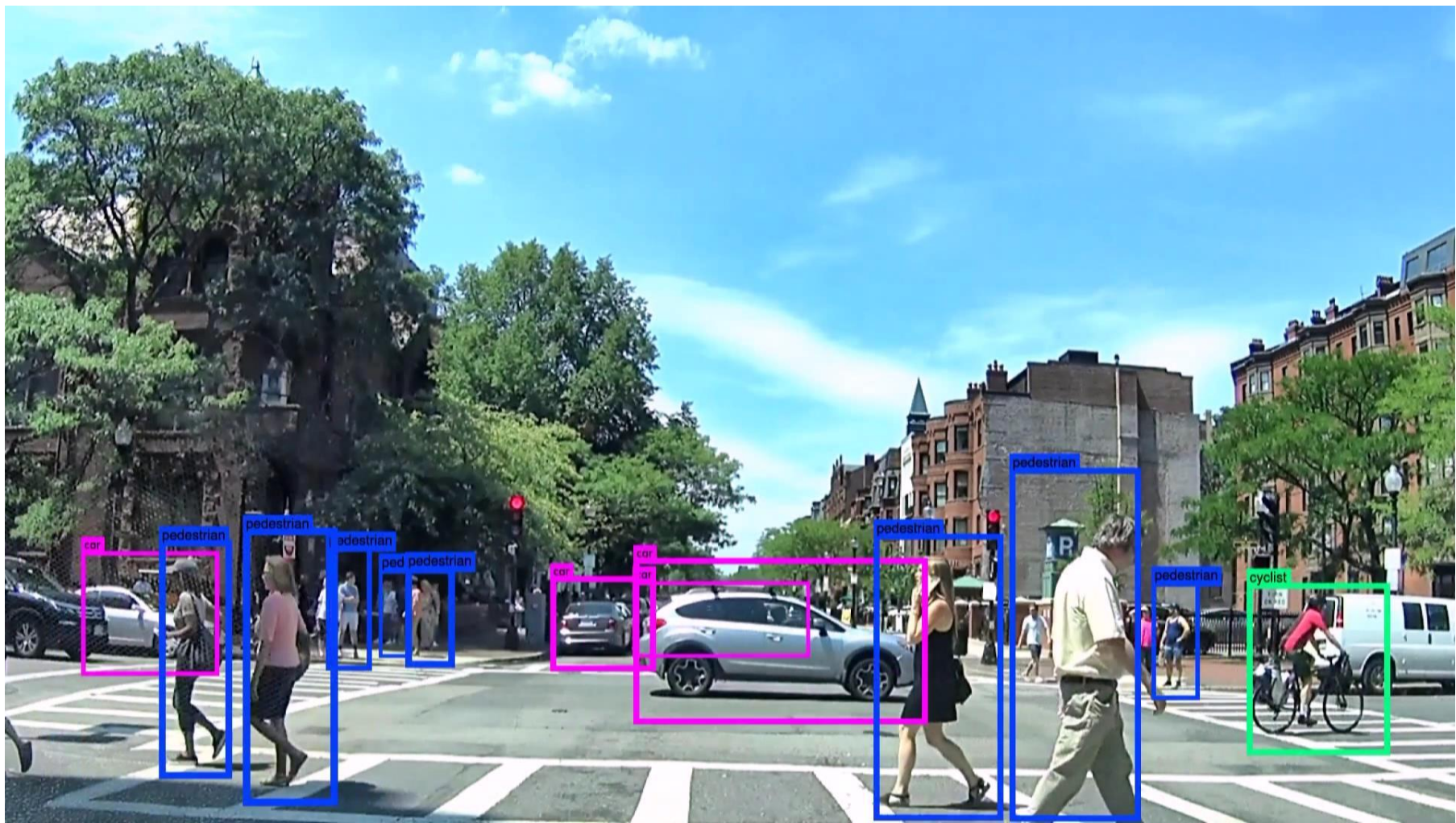
- Census responses are data
- Blueprints are data



What is data?

Data is a representation of objects or information from the real world

- Census responses are data
- Blueprints are data
- Images are data



What is data?

Data is a representation of objects or information from the real world

- Census responses are data
- Blueprints are data
- Images are data
- Financial records are data

Cash Flow

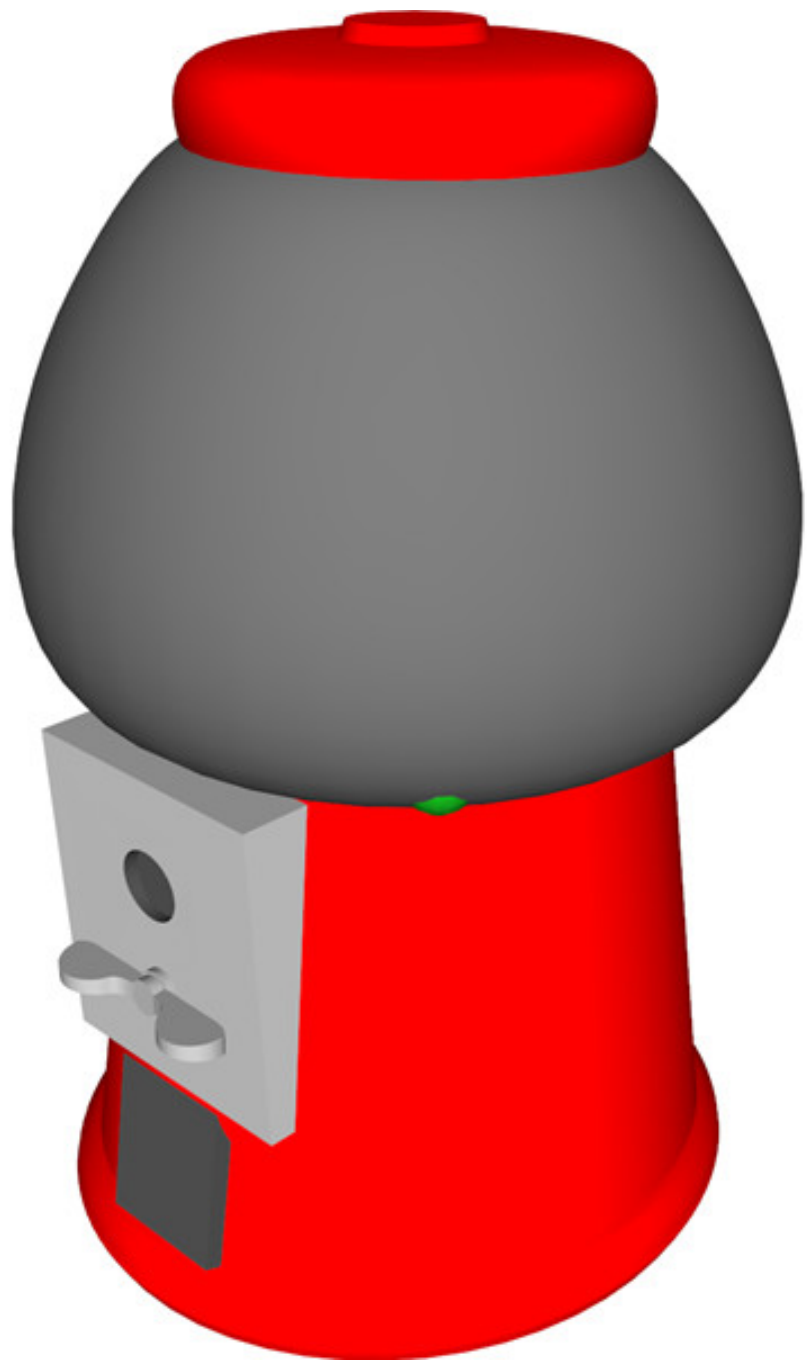
Date: February 2013

Account	Description	March 2000	February 2000	Current Change	YTD March 2000
Cash Flow From Operations					
0	Net Income	365,102	361,218	3,884	3,884,060
1100	Net change in Accounts receivable - trade	202,826	(299,201)	502,028	(3,067,762)
1110	Net change in Accounts receivable - other	1,310	1,179	131	14,938
1199	Net change in Allowance for bad debts	4,373	4,327	47	46,523
1200	Net change in Inventory	(5,386)	(10,027)	4,642	(208,522)
1900	Net change in Other current assets	(1,804)	(2,035)	231	(22,825)
3000	Net change in Accounts payable	40,428	40,086	342	595,083
3100	Net change in Accrued payroll	(491)	(406)	(85)	(5,492)
3400	Net change in Other current liabilities	12,819	12,146	673	141,702
Total	Cash Flow From Operations	619,177	107,286	511,891	1,377,705
Cash Flow From Investments					
2900	Increases in Other assets	(398)	(420)	22	(5,126)
2000	Decreases in Fixed assets	354	346	8	4,027
2100	Decreases in Accumulated depreciation	23,107	22,862	246	245,823
Total	Cash Flow From Investments	23,063	22,788	276	244,724
Cash Flow From Financing					
3500	Increases in Long-term notes	279	318	(39)	4,163
3900	Increases in Other liabilities	261	189	72	2,506
4000	Increases in Capital stock	211	203	8	2,501
Total	Cash Flow From Financing	751	710	41	9,170
Total Cash Flow		642,991	130,783	512,208	1,631,599

What is data?

Data is a representation of objects or information from the real world

- Census responses are data
- Blueprints are data
- Images are data
- Financial records are data
- CAD designs are data



Defining Data Analytics

Good luck! 😜

But seriously, **Data Analytics** is a notoriously broad term describing pretty much anything that uses some data with any amount of analysis of that data to make decisions. Let's try some better terminology...

Definitions

Predictive Analytics - Using data to predict outcomes on an individual level

Forecasting - Using data to predict macro-level outcomes

Why now? Why not when computers were invented?

- More Data
 - There has been a tremendous increase in the amount of data available to companies
- More Computer Power
 - In 2005, my laptop had $\frac{3}{4}$ GB of RAM, and cost over \$2000
 - As of 2014, my laptop (still at just over \$2000) had 16GB of RAM, nearly a 2000% increase! The same was true of the increase in processor performance.
 - Today, a similar laptop can be purchased for closer to \$1000

Why now?

- Synergy between Data and Computation Power
 - We now have enough data to make meaningful conclusions
 - We have enough computational power!
 - [OmniSci](#), [more OmniSci](#)
 - NVIDIA GTX 1080 Ti: ~11 TFlops (11 trillion operations per second)
 - Only costs \$599

So where can we get the data we need?

Anywhere that records are kept, if the owner is willing to share.

- Facebook, Twitter, and LinkedIn all collect very robust information on users and their interactions
- We could also record our own data if needed: surveys, usage statistics, credit card transactions, etc.

So where can we get the data we need?

Some of my all-time favorites:

- IPUMS.org
- Data.gov
- ICPSR@UMich

So the data will magically analyze itself, right?!

- Just like a book that is never read, data that is only recorded and never used is not that useful
- We need to find a way to UTILIZE data
- That is what we will learn about during the rest of this class!

We need to ask questions!

Exploring data helps us to find the questions we care about

Let's use Tableau to explore our data

For lab today

- Complete part 1 of the first project
 - Find a group (use Slack!)
 - Play with the [example data](#) in Tableau
 - Variable descriptions are available [here](#)
 - Choose a research question

If you need, remember that Tableau provides all sorts of great training videos. I have posted some of my favorites on Canvas.