

## What is a Data Model?

## Hands-On

**Reading:** Introduction to CSV Data  
10 min

**Video:** Introduction to CSV Data  
4 min

## Different Kinds of Data Models (Part 1)

**Video:** What is a Relational Data Model?  
10 min

**Reading:** Slides: What is A Relational Data Model?  
10 min

**Video:** What is a Semistructured Data Model?  
6 min

**Reading:** Slides: What is a Semistructured Data Model?  
10 min

**Discussion Prompt:**  
Let's discuss: Utilization of XML or JSON on the Internet  
5 min

## Hands-On

By the end of this activity, you will be able to:

1. Identify the key features in CSV data
2. Import CSV data to a spreadsheet and plot values

Step 1. **Open a terminal shell.** Open a terminal shell by clicking on the square black box on the top left of the screen.



Run `cd Downloads/big-data-2/csv` to change into the directory containing the csv file. (This was downloaded in Week 1 <https://www.coursera.org/learn/big-data-management/supplement/YVDP/instructions-for-downloading-hands-on-datasets>)

```
1 cd Downloads/big-data-2/csv
```

Step 2. **Look at CSV file.** The CSV file contains census data for the United States. Run `ls` to see the name of the CSV file.

```
1 ls
```

```
[cloudera@quickstart ~]$ cd Downloads/big-data-2/csv
[cloudera@quickstart csv]$ ls
census.csv
```

Run `more census.csv` to look at the contents of the CSV file.

```
1 more census.csv
```

The first line of the file is the head and the remaining lines are the data. Each entry in the file is separated by a comma.

Entry, or Column

Header

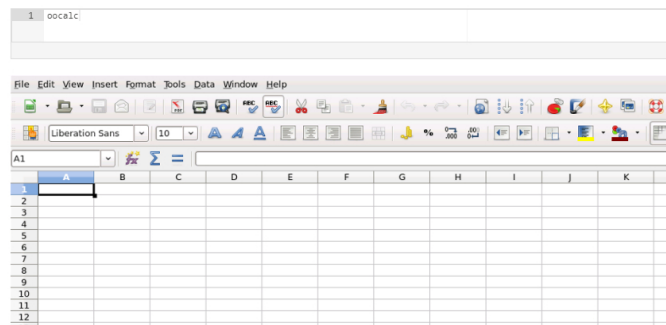
Data

```
N,DIVISION,STATE,COUNTY,STNAME,CTYNAM,CENSUS2010POP,ESTIMATESBASE2010,POPESTIMATE2010,POPESTIMATE2011,POPESTIMATE2012,POPESTIMATE2013,POPESTIMATE2014,POPESTIMATE2015,NPOPCHG_2010,NPOPCHG_2011,NPOPCHG_2012,NPOPCHG_2013,NPOPCHG_2014,NPOPCHG_2015,BIRTHS2010,BIRTHS2011,BIRTHS2012,BIRTHS2013,BIRTHS2014,BIRTHS2015,DEATHS2010,DEATHS2011,DEATHS2012,DEATHS2013,DEATHS2014,DEATHS2015,NATURALINC2010,NATURALINC2011,NATURALINC2012,NATURALINC2013,NATURALINC2014,NATURALINC2015,INTERNATIONALMIG2010,INTERNATIONALMIG2011,INTERNATIONALMIG2012,INTERNATIONALMIG2013,INTERNATIONALMIG2014,INTERNATIONALMIG2015,DOMESTICMIG2010,DOMESTICMIG2011,DOMESTICMIG2012,DOMESTICMIG2013,DOMESTICMIG2014,DOMESTICMIG2015,NETMIG2010,NETMIG2011,NETMIG2012,NETMIG2013,NETMIG2014,NETMIG2015,RESTOTAL2010,RESTOTAL2011,RESTOTAL2012,RESTOTAL2013,RESTOTAL2014,RESTOTAL2015,STIMATES2010,GUESTIMATES2011,GUESTIMATES2012,GUESTIMATES2013,GUESTIMATES2014,GUESTIMATES2015,RTHCH2010,RTHCH2011,RTHCH2012,RTHCH2013,RTHCH2014,RTHCH2015,RDEATH2010,RDEATH2011,RDEATH2012,RDEATH2013,RDEATH2014,RDEATH2015,RNATURALINC2010,RNATURALINC2011,RNATURALINC2012,RNATURALINC2013,RNATURALINC2014,RNATURALINC2015,RINTERNATIONALMIG2010,RINTERNATIONALMIG2011,RINTERNATIONALMIG2012,RINTERNATIONALMIG2013,RINTERNATIONALMIG2014,RINTERNATIONALMIG2015,RDOMESTICMIG2010,RDOMESTICMIG2011,RDOMESTICMIG2012,RDOMESTICMIG2013,RDOMESTICMIG2014,RDOMESTICMIG2015,RNETMIG2010,RNETMIG2011,RNETMIG2012,RNETMIG2013,RNETMIG2014,RNETMIG2015
```

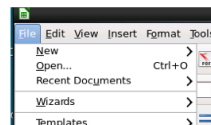
```
00,Alabama,Alabama,4779736,4780127,4785161,4881109,4816089,4839533,4846411,4858979,4876125,4881226,4886266,4891308,4896350,4901392,4906434,4911476,4916518,4921560,4926602,4931644,4936686,4941728,4946770,4951812,4956854,4961896,4966938,4971980,4977022,4982064,4987106,4992148,4997190,5002232,5007274,5012316,5017358,5022400,5027442,5032484,5037526,5042568,5047610,5052652,5057694,5062736,5067778,5072820,5077862,5082904,5087946,5092988,5098030,5103072,5108114,5113156,5118198,5123240,5128282,5133324,5138366,5143408,5148450,5153492,5158534,5163576,5168618,5173660,5178702,5183744,5188786,5193828,5198870,5203912,5208954,5213996,5219038,5224080,5229122,5234164,5239206,5244248,5249290,5254332,5259374,5264416,5269458,5274500,5279542,5284584,5289626,5294668,5299710,5304752,5309794,5314836,5319878,5324920,5329962,5335004,5340046,5345088,5350130,5355172,5360214,5365256,5370298,5375340,5380382,5385424,5390466,5395508,5400550,5405592,5410634,5415676,5420718,5425760,5430802,5435844,5440886,5445928,5450970,5456012,5461054,5466096,5471138,5476180,5481222,5486264,5491306,5496348,5501390,5506432,5511474,5516516,5521558,5526600,5531642,5536684,5541726,5546768,5551810,5556852,5561894,5566936,5571978,5577020,5582062,5587104,5592146,5597188,5602230,5607272,5612314,5617356,5622398,5627440,5632482,5637524,5642566,5647608,5652650,5657692,5662734,5667776,5672818,5677860,5682902,5687944,5692986,5698028,5703070,5708112,5713154,5718196,5723238,5728280,5733322,5738364,5743406,5748448,5753490,5758532,5763574,5768616,5773658,5778700,5783742,5788784,5793826,5798868,5803910,5808952,5813994,5819036,5824078,5829120,5834162,5839204,5844246,5849288,5854330,5859372,5864414,5869456,5874498,5879540,5884582,5889624,5894666,5899708,5904750,5909792,5914834,5919876,5924918,5929960,5934994,5940036,5945078,5950120,5955162,5960204,5965246,5970288,5975330,5980372,5985414,5990456,5995498,6000540,6005582,6010624,6015666,6020708,6025750,6030792,6035834,6040876,6045918,6050960,6056002,6061044,6066086,6071128,6076170,6081212,6086254,6091296,6096338,6101380,6106422,6111464,6116506,6121548,6126590,6131632,6136674,6141716,6146758,6151800,6156842,6161884,6166926,6171968,6177010,6182052,6187094,6192136,6197178,6202220,6207262,6212304,6217346,6222388,6227430,6232472,6237514,6242556,6247598,6252640,6257682,6262724,6267766,6272808,6277850,6282892,6287934,6292976,6298018,6303060,6308102,6313144,6318186,6323228,6328270,6333312,6338354,6343396,6348438,6353480,6358522,6363564,6368606,6373648,6378690,6383732,6388774,6393816,6398858,6403900,6408942,6413984,6419026,6424068,6429110,6434152,6439194,6444236,6449278,6454320,6459362,6464404,6469446,6474488,6479530,6484572,6489614,6494656,6499698,6504740,6509782,6514824,6519866,6524908,6529950,6534992,6540034,6545076,6550118,6555160,6560202,6565244,6570286,6575328,6580370,6585412,6590454,6595496,6600538,6605580,6610622,6615664,6620706,6625748,6630790,6635832,6640874,6645916,6650958,6656000,6661042,6666084,6671126,6676168,6681210,6686252,6691294,6696336,6701378,6706420,6711462,6716504,6721546,6726588,6731630,6736672,6741714,6746756,6751798,6756840,6761882,6766924,6771966,6777008,6782050,6787092,6792134,6797176,6802218,6807260,6812302,6817344,6822386,6827428,6832470,6837512,6842554,6847596,6852638,6857680,6862722,6867764,6872806,6877848,6882890,6887932,6892974,6898016,6903058,6908100,6913142,6918184,6923226,6928268,6933310,6938352,6943394,6948436,6953478,6958520,6963562,6968604,6973646,6978688,6983730,6988772,6993814,6998856,7003898,7008940,7013982,7019024,7024066,7029108,7034150,7039192,7044234,7049276,7054318,7059360,7064402,7069444,7074486,7079528,7084570,7089612,7094654,7099696,7104738,7109780,7114822,7119864,7124906,7129948,7134990,7140032,7145074,7150116,7155158,7160200,7165242,7170284,7175326,7180368,7185410,7190452,7195494,7200536,7205578,7210620,7215662,7220704,7225746,7230788,7235830,7240872,7245914,7250956,7255998,7261040,7266082,7271124,7276166,7281208,7286250,7291292,7296334,7301376,7306418,7311460,7316502,7321544,7326586,7331628,7336670,7341712,7346754,7351796,7356838,7361880,7366922,7371964,7377006,7382048,7387090,7392132,7397174,7402216,7407258,7412300,7417342,7422384,7427426,7432468,7437510,7442552,7447594,7452636,7457678,7462720,7467762,7472804,7477846,7482888,7487930,7492972,7498014,7503056,7508098,7513140,7518182,7523224,7528266,7533308,7538350,7543392,7548434,7553476,7558518,7563560,7568602,7573644,7578686,7583728,7588770,7593812,7598854,7603896,7608938,7613980,7619022,7624064,7629106,7634148,7639190,7644232,7649274,7654316,7659358,7664400,7669442,7674484,7679526,7684568,7689610,7694652,7699694,7704736,7709778,7714820,7719862,7724904,7729946,7734988,7740030,7745072,7750114,7755156,7760198,7765240,7770282,7775324,7780366,7785408,7790450,7795492,7800534,7805576,7810618,7815660,7820702,7825744,7830786,7835828,7840870,7845912,7850954,7855996,7861038,7866080,7871122,7876164,7881206,7886248,7891290,7896332,7901374,7906416,7911458,7916500,7921542,7926584,7931626,7936668,7941710,7946752,7951794,7956836,7961878,7966920,7971962,7977004,7982046,7987088,7992130,7997172,8002214,8007256,8012298,8017340,8022382,8027424,8032466,8037508,8042550,8047592,8052634,8057676,8062718,8067760,8072802,8077844,8082886,8087928,8092970,8098012,8103054,8108096,8113138,8118180,8123222,8128264,8133306,8138348,8143390,8148432,8153474,8158516,8163558,8168600,8173642,8178684,8183726,8188768,8193810,8198852,8203894,8208936,8213978,8219020,8224062,8229104,8234146,8239188,8244230,8249272,8254314,8259356,8264398,8269440,8274482,8279524,8284566,8289608,8294650,8299692,8304734,8309776,8314818,8319860,8324902,8329944,8334986,8340028,8345070,8350112,8355154,8360196,8365238,8370280,8375322,8380364,8385406,8390448,8395490,8400532,8405574,8410616,8415658,8420700,8425742,8430784,8435826,8440868,8445910,8450952,8455994,8461036,8466078,8471120,8476162,8481204,8486246,8491288,8496330,8501372,8506414,8511456,8516498,8521540,8526582,8531624,8536666,8541708,8546750,8551792,8556834,8561876,8566918,8571960,8577002,8582044,8587086,8592128,8597170,8602212,8607254,8612296,8617338,8622380,8627422,8632464,8637506,8642548,8647590,8652632,8657674,8662716,8667758,8672800,8677842,8682884,8687926,8692968,8698010,8703052,8708094,8713136,8718178,8723220,8728262,8733304,8738346,8743388,8748430,8753472,8758514,8763556,8768598,8773640,8778682,8783724,8788766,8793808,8798850,8803892,8808934,8813976,8819018,8824060,8829102,8834144,8839186,8844228,8849270,8854312,8859354,8864396,8869438,8874480,8879522,8884564,8889606,8894648,8899690,8904732,8909774,8914816,8919858,8924900,8929942,8934984,8940026,8945068,8950110,8955152,8960194,8965236,8970278,8975320,8980362,8985404,8990446,8995488,9000530,9005572,9010614,9015656,9020698,9025740,9030782,9035824,9040866,9045908,9050950,9055992,9061034,9066076,9071118,9076160,9081202,9086244,9091286,9096328,9101370,9106412,9111454,9116496,9121538,9126580,9131622,9136664,9141706,9146748,9151790,9156832,9161874,9166916,9171958,9177000,9182042,9187084,9192126,9197168,9202210,9207252,9212294,9217336,9222378,9227420,9232462,9237504,9242546,9247588,9252630,9257672,9262714,9267756,9272798,9277840,9282882,9287924,9292966,9298008,9303050,9308092,9313134,9318176,9323218,9328260,9333302,9338344,9343386,9348428,9353470,9358512,9363554,9368596,9373638,9378680,9383722,9388764,9393806,9398848,9403890,9408932,9413974,9419016,9424058,9429100,9434142,9439184,9444226,9449268,9454310,9459352,9464394,9469436,9474478,9479520,9484562,9489604,9494646,9499688,9504730,9509772,9514814,9519856,9524898,9529940,9534982,9540024,9545066,9550108,9555150,9560192,9565234,9570276,9575318,9580360,9585402,9590444,9595486,9600528,9605570,9610612,9615654,9620696,9625738,9630780,9635822,9640864,9645906,9650948,9655990,9661032,9666074,9671116,9676158,9681200,9686242,9691284,9696326,9701368,9706410,9711452,9716494,9721536,9726578,9731620,9736662,9741704,9746746,9751788,9756830,9761872,9766914,9771956,9777000,9782042,9787084,9792126,9797168,9802210,9807252,9812294,9817336,9822378,9827420,9832462,9837504,9842546,9847588,9852630,9857672,9862714,9867756,9872798,9877840,9882882,9887924,9892966,9898008,9903050,9908092,9913134,9918176,9923218,9928260,9933302,9938344,9943386,9948428,9953470,9958512,9963554,9968596,9973638,9978680,9983722,9988764,9993806,9998848,10000000
```

Hit the spacebar to scroll down, and `q` to quit more.

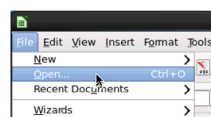
Step 3. **Open spreadsheet application.** Run `oocalc` to start the spreadsheet application.



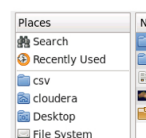
Step 4. **Import CSV to spreadsheet.** Let's import the CSV file to the spreadsheet by clicking on *File*:

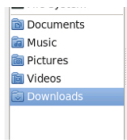


Next, click *Open*:

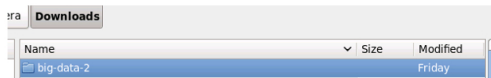


Next, click *Downloads* in the Places panel:

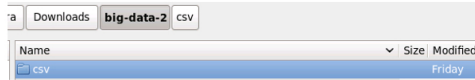




Next, double-click *big-data-2* in the file pane:



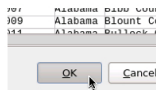
Next, double-click *csv*:



Next, double-click *census.csv*:



In the Text Import dialog, click *OK*:



The CSV data is now loaded into the spreadsheet.

	A	B	C	D	E	F	G	H	I
1	SUMLEV	REGION	DIVISION	STATE	COUNTY	STNAME	CTYNAME	CENSUS2010POP	EST
2	40	3	6	1	0	Alabama	Alabama	4779736	
3	50	3	6	1	1	Alabama	Autauga County	54571	
4	50	3	6	1	3	Alabama	Baldwin County	182265	
5	50	3	6	1	5	Alabama	Barbour County	27457	
6	50	3	6	1	7	Alabama	Bibb County	22915	
7	50	3	6	1	9	Alabama	Blount County	57322	
8	50	3	6	1	11	Alabama	Bullock County	10914	
9	50	3	6	1	13	Alabama	Butler County	20947	
10	50	3	6	1	15	Alabama	Calhoun County	118572	
11	50	3	6	1	17	Alabama	Chambers County	34215	
12	50	3	6	1	19	Alabama	Cherokee County	25689	

Step 5. **See size of CSV.** Scroll to the bottom of the spreadsheet to see the size of the CSV file.

3191	50	4	8	56	39	Wyoming	Leton County	21294	
3192	50	4	8	56	41	Wyoming	Linton County	21118	
3193	50	4	8	56	43	Wyoming	Washakie County	8533	
3194	50	4	8	56	45	Wyoming	Weston County	7208	
3195									

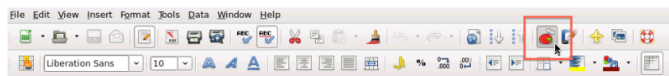
There are 3194 rows. If the CSV file had millions or more rows, then we could not import it into a spreadsheet. In this case, we would need a Big Data system such as Hadoop to analyze the data.

Scroll back to the top.

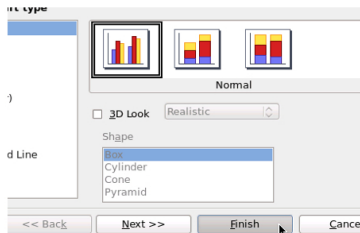
Step 6. **Create chart.** Let's create a chart of the estimated population of the state of Alabama. Row 2 contains the data for Alabama. Select cells in row 2 and columns J through O to get the estimated population for 2010 through 2015.

	J	K	L	M	N	O
1	POPESTIMATE2010	POPESTIMATE2011	POPESTIMATE2012	POPESTIMATE2013	POPESTIMATE2014	POPESTIMATE2015
2	4785161	4801108	4816089	4830533	4846411	4858979

Click on the chart button:



Click *Finish* to display the chart:



The chart should be displayed in the spreadsheet:

