

Analyzing the FDIC DataSet

Wolfram has expanded its Knowledgebase of data resources considerably, adding several rich and interesting datasets.

The Federal Deposit Insurance Corporation (FDIC) is an independent US government agency charged with insuring deposits in US financial institutions. Presently, deposits in member institutions are insured up to \$250, 000. Here, the holdings, size distribution, and geographic dispersal of member institutions are explored.

The FDIC

We begin by loading the FDIC dataset as a resource:

```
fdic =  
  ResourceData[ResourceObject[Association["Name" → "FDIC Institution EntityStore",  
    "UUID" → "6f5d37d4-1406-483c-b67c-f58d903d16b1",  
    "ResourceType" → "DataResource", "Version" → "1.0.0",  
    "Description" → "A Wolfram Language EntityStore with selected  
      \ data on FDIC insured institutions",  
    "ContentSize" → Quantity[0, "Bytes"], "ContentElements" → {"EntityStore"}]]]
```

```
EntityStore[ [ Type: FDICInstitution  
  Entities: {10002, 10004, 10007, ..., 999, 9997, 9998} (6131)  
  Properties: {AdjustedAverageAssetsForLeverage, AdjustedTotalRiskWeightedAssets,  
    AllOtherAssets, ..., UnusedLoanCommitments, VolatileLiabilities, ZIPCode} (54) ] ]
```

```
PrependTo[$EntityStores, fdic];
```

There are over six thousand entities listed in the dataset:

```
Length[ents = EntityList["FDICInstitution"]]
```

```
6131
```

The dataset contains a large number of properties, for example:

```
In[109]:= Take[EntityProperties["FDICInstitution"] // Sort, 20]
```

```
Out[109]= { adjusted average assets for leverage , adjusted total risk weighted assets ,  
  all other assets , all other liabilities , average total assets , bank equity ,  
  bank premises and fixed assets , cash and balances due from depository Institutions ,  
  city , common stock , county , domestic deposits , earning assets ,  
  FDIC certificate number , federal reserve ID number , interest bearing balances ,  
  interest bearing deposits , label , loan loss allowance , long term assets }
```

We can, for example, plot the geolocation of a selection of entities:

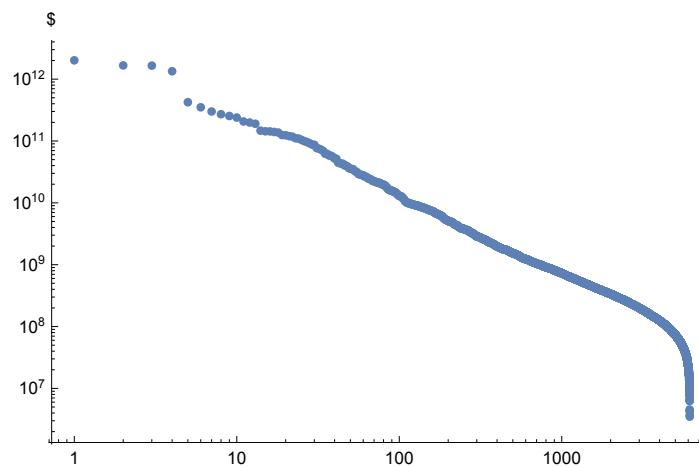
```
GeoListPlot[Take[EntityList["FDICInstitution"], 50],  
 PlotMarkers -> "$", ImageSize -> Medium]
```



Assets

The dataset shows the enormous size of a handful of the largest financial institutions:

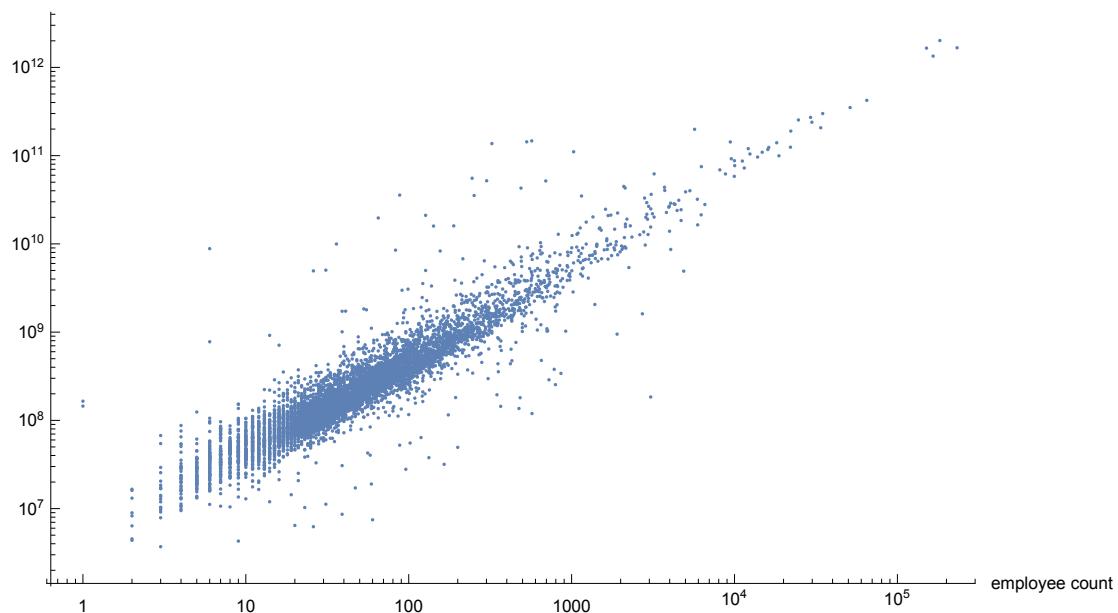
```
ListLogLogPlot[Reverse@Sort[EntityValue["FDICInstitution", "TotalAssets"]],  
 AxesLabel -> Automatic, PlotStyle -> PointSize[Medium]]
```



There is a roughly linear relationship between total assets and employee count, although with considerable variation in productivity (as measured by assets per employee) amongst the smaller institutions:

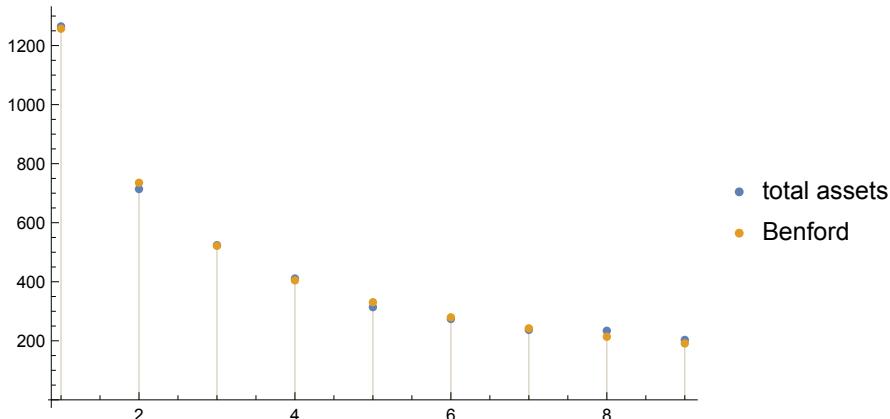
```
empVsAssets =
EntityValue["FDICInstitution", {"TotalEmployeeNumber", "TotalAssets"}];
ListLogLogPlot[empVsAssets, AxesLabel -> {"employee count", "total assets"}, ImageSize -> Large]
```

total assets



The distribution of assets per city fulfills Benford's law to a remarkable degree:

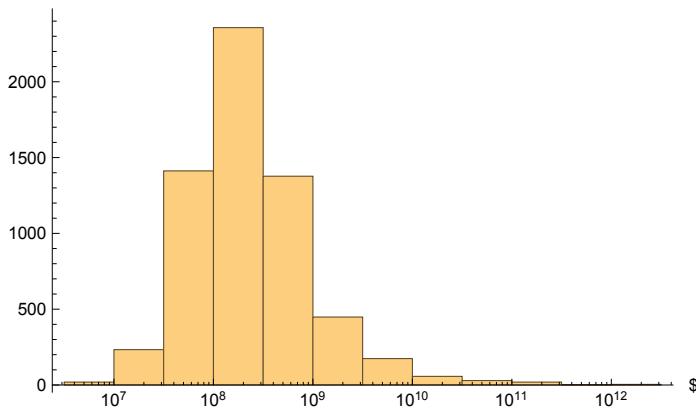
```
$percity = Reverse[SortBy[{#[[1, 1]], Total#[[All, -1]]}] & /@
Normal[Normal[DeleteMissing[SortBy[GroupBy[EntityValue["FDICInstitution",
{"City", "TotalAssets"}], First], Length], 1, 3]][[All, -1]]], Last]];
ListPlot[{Tally[IntegerDigits[Round[QuantityMagnitude[#[2]]]]][[1]] &@@@ $percity],
Table[{d, Length[$percity] Log10[1 + 1/d]}, {d, 1, 9}],
PlotRange -> All, Filling -> Axis, PlotLegends -> {"total assets", "Benford"}]
```



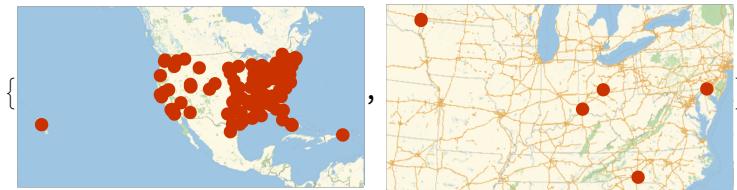
The Big Six

A small handful of banks control a disproportionately large share of total assets:

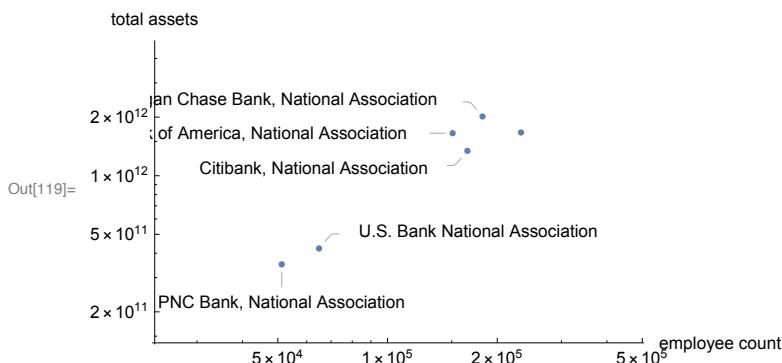
```
dat = EntityValue["FDICInstitution", "TotalAssets", "EntityAssociation"];
Histogram[dat, "Log", AxesLabel → Automatic]
```



```
GeoListPlot[Keys[Select[dat, GreaterThan[Quantity[#, "USDollars"]]]]] & /@
{5*^9, 300*^9}]
```



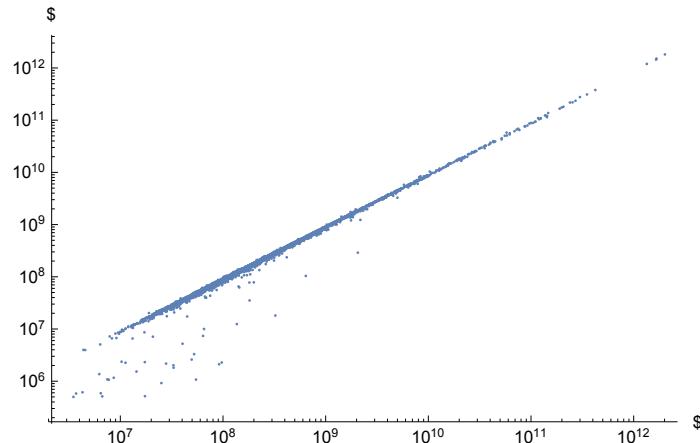
```
In[119]:= ListLogLogPlot[Callout[{#2, #3}, #1] & @@ EntityValue[
EntityList[EntityClass["FDICInstitution", "TotalAssets" → TakeLargest[6]]],
 {"Label", "TotalEmployeeNumber", "TotalAssets"}], PlotRangePadding → Scaled[0.25], AxesLabel → {"employee count", "total assets"}, ImageSize → Medium]
```



Assets and Liabilities

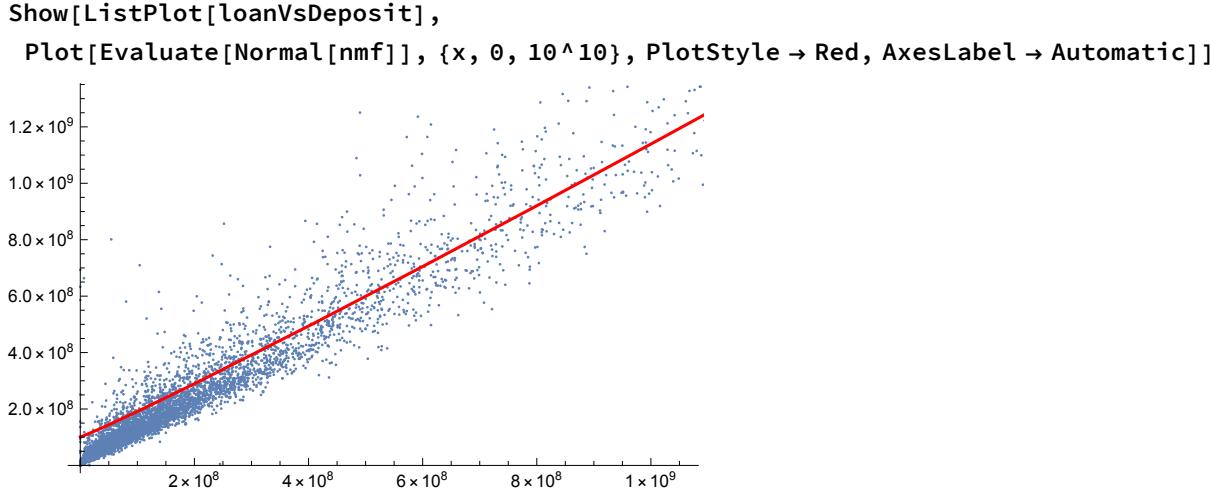
As expected, there is a log-linear relationship between total assets and liabilities:

```
assetsVsLiability =
  EntityValue["FDICInstitution", {"TotalAssets", "TotalLiabilities"}];
ListLogLogPlot[assetsVsLiability, AxesLabel → Automatic]
```



The same is also true for the relationship between loans and deposits:

```
loanVsDeposit =
  EntityValue["FDICInstitution", {"NetLoansAndLeases", "TotalDeposits"}];
nmf = NonlinearModelFit[Select[QuantityMagnitude /@ loanVsDeposit, Min[#] > 0 &],
  c + a x^α, {a, α, c}, x]
FittedModel[ 1.00786×10^8+0.298151 x1.06021 ]
```



```

props = {EntityProperty["FDICInstitution", "TotalAssets"],
  EntityProperty["FDICInstitution", "LongTermAssets"],
  EntityProperty["FDICInstitution", "TotalLiabilities"]};
data = Select[N@QuantityMagnitude@DeleteMissing[EntityValue[
  "FDICInstitution", Append[props, "Entity"]]], 1, 2], Min[Most[#]] > 0 &];

data2 = Select[{#1/#2, #1/#3, #4} &@@@ data,
  Between[{0, 15}]#[[1]]] && Between[{1, 1.4}]#[[2]]] &;

sdk = SmoothKernelDistribution[Most /@ data2];
label[{a_, b_, c_, _d_}] := Column[{d, Grid[Transpose[{props,
  Quantity[NumberForm[#, 4], "USDollars"] &/@{a, b, c}}]], Dividers -> Center}]}

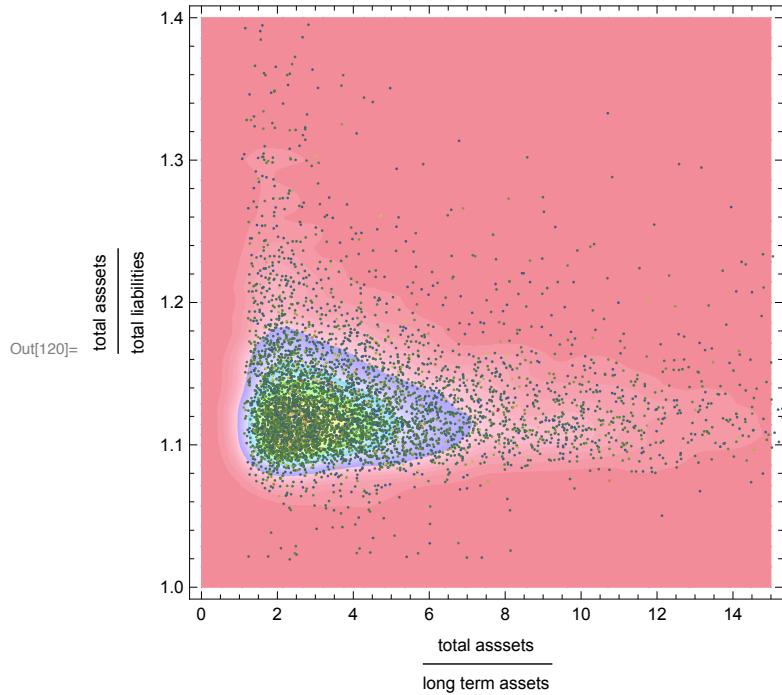
```

```

In[120]:= ContourPlot[PDF[sdk, {x, y}], {x, 0, 15}, {y, 1, 1.4},
 PlotRange -> All, Contours -> 50, ContourStyle -> None,
 ColorFunction -> (Lighter[ColorData["BrightBands"]#[], 0.5] &), FrameLabel ->
 {"total assets"/"long term assets", "total assets"/"total liabilities"},

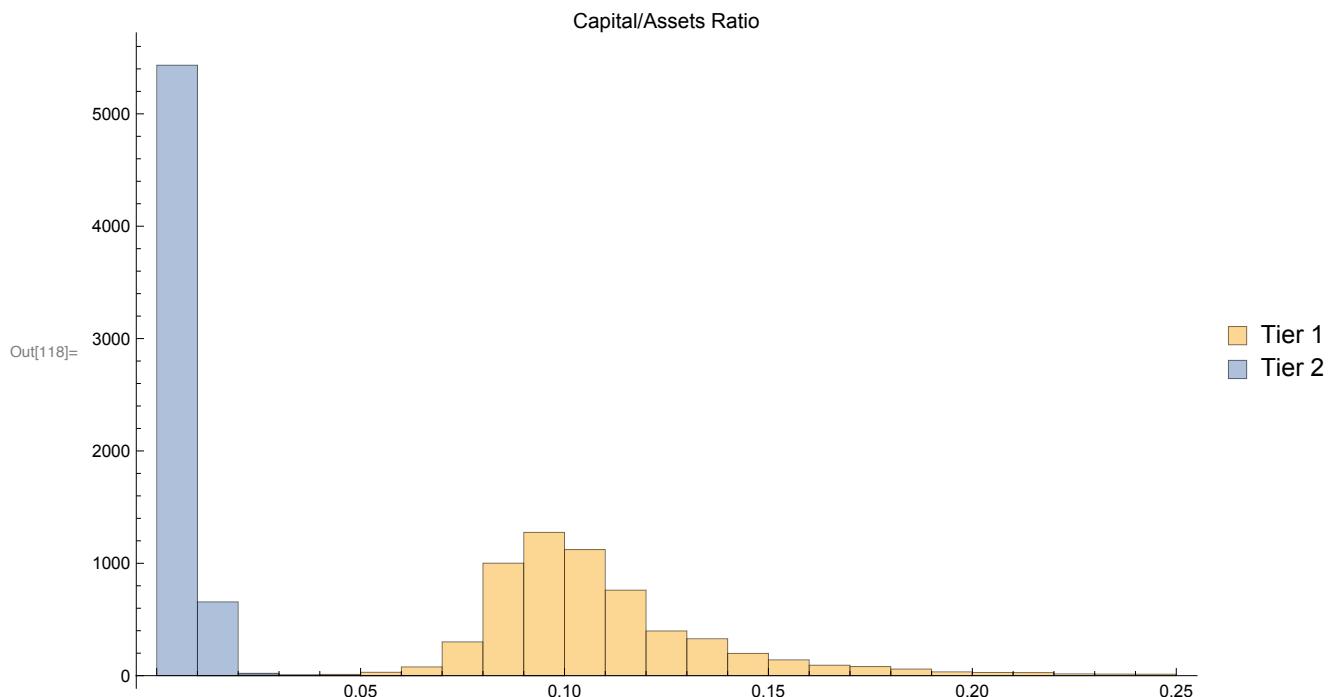
Epilog -> ({ColorData["DarkRainbow"][(Log10[#1] - 7)/5], PointSize[0.00125],
  Point[{#1/#2, #1/#3}]} &@@@ data), ImageSize -> Medium]

```



Tier 1 and Tier 2 Capital

```
In[117]:= {tierOneCapitalToAssets, tierTwoCapitalToAssetsCapitalToAssets} =
  (Divide @@ EntityValue["FDICInstitution", {#, "TotalAssets"}]) & /@
  {"TierOneCapital", "TierTwoRiskBasedCapital"};
Histogram[{tierOneCapitalToAssets, tierTwoCapitalToAssetsCapitalToAssets},
{0, 0.25, 0.01}, ChartLegends -> {"Tier 1", "Tier 2"},
PlotLabel -> "Capital/Assets Ratio", ImageSize -> Large]
```



Deposits

We can plot the location of the largest deposit-taking institutions:

```
dat1 = EntityValue["FDICInstitution", "TotalDeposits", "EntityAssociation"];
{bottom1percent, top1percent} = Quantile[values = Values[dat1], {0.01, 0.99}]
{ $11 361 000 , $20 683 779 000 }
```

```
GeoListPlot[Keys[Select[data1, GreaterThan[top1percent]]]]
```

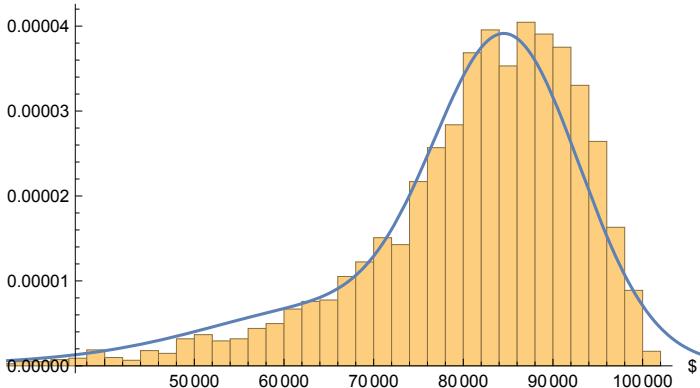


The distribution of insured deposits is heavily skewed and is best modelled by a Gaussian-Logistic mixture distribution:

```
H1 = Histogram[
  PID = EntityValue["FDICInstitution", percent insured deposits, "EntityAssociation"],
  40, "PDF", AxesLabel → Automatic];

PIDDistn = FindDistribution[QuantityMagnitude@Values@PID]
MixtureDistribution[{0.249604, 0.750396},
 {LogisticDistribution[63909.1, 9322.17], NormalDistribution[84889.4, 8136.52]}]

Show[{H1, Plot[PDF[PIDDistn, x], {x, 0, 120000}]}]
```



Financial Leverage

The most highly leveraged institutions:

```
dataSet3 =
Append[#, <|"FinancialLeverage" → N[#TotalLiabilities / #TotalEquity]|>] & /@
EntityValue["FDICInstitution", {"TotalLiabilities", "TotalEquity"}, "Dataset"];
dataSet3[TakeLargestBy["FinancialLeverage", 10]]
```

	TotalLiabilities	TotalEquity	FinancialLever
Park Federal Savings Bank	\$137631000	\$2694000	51.088
First CornerStone Bank	\$101202000	\$2105000	48.077
American Patriot Bank	\$61301000	\$1277000	48.0039
Foothills Community Bank	\$68208000	\$1455000	46.8784
Pinnacle Bank	\$123789000	\$2671000	46.3456
Illinois–Service Federal Savings and Loan Association	\$98867000	\$2137000	46.2644
Harvest Community Bank	\$135673000	\$3118000	43.5128
Cecil Bank	\$247296000	\$5893000	41.9644
Guaranty Bank	\$1001119000	\$23958000	41.7864
Horry County State Bank	\$354900000	\$8515000	41.6794

```
Histogram[dataSet3[All, "FinancialLeverage"], {0, 20, 0.5},
PlotLabel → "FDIC banks Financial Leverage distribution"]
```

