k-means Clustering

Theme

Initialization

In[13]:=
```
data = {{0, 0}, {1, 0}, {1.5, 0.8}, {1.5, 1}};
dataLabels = Subscript[bi["p"], #] & /@ Range[Length[data]];
cluster = {{1.35, 0.7}, {1.6, 1.1}};
```

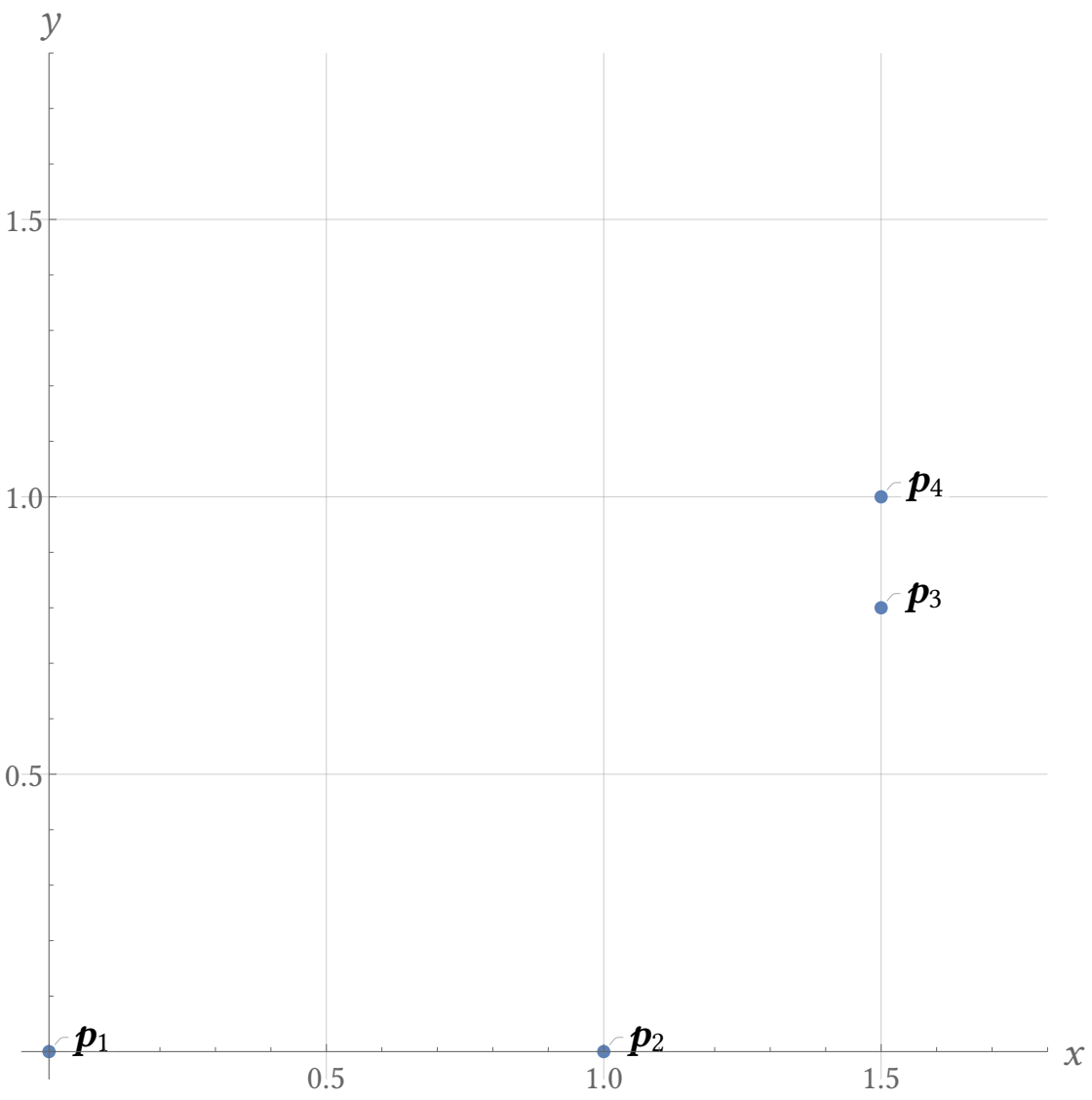
In[16]:=
```
assignPointsToClusters[] := Block[{},
  cluster1 = {};
  cluster2 = {};
  Do[
    { AppendTo[cluster1, data[[i]]]  Nearest[cluster, data[[i]]][[1]] == cluster[[1]]
    { AppendTo[cluster2, data[[i]]]  Nearest[cluster, data[[i]]][[1]] == cluster[[2]]
   , {i, 1, Length[data]}]
 ]

calculateClusters[] := Block[{},
  cluster[[1]] = Mean[cluster1];
  cluster[[2]] = Mean[cluster2];
 ]
```

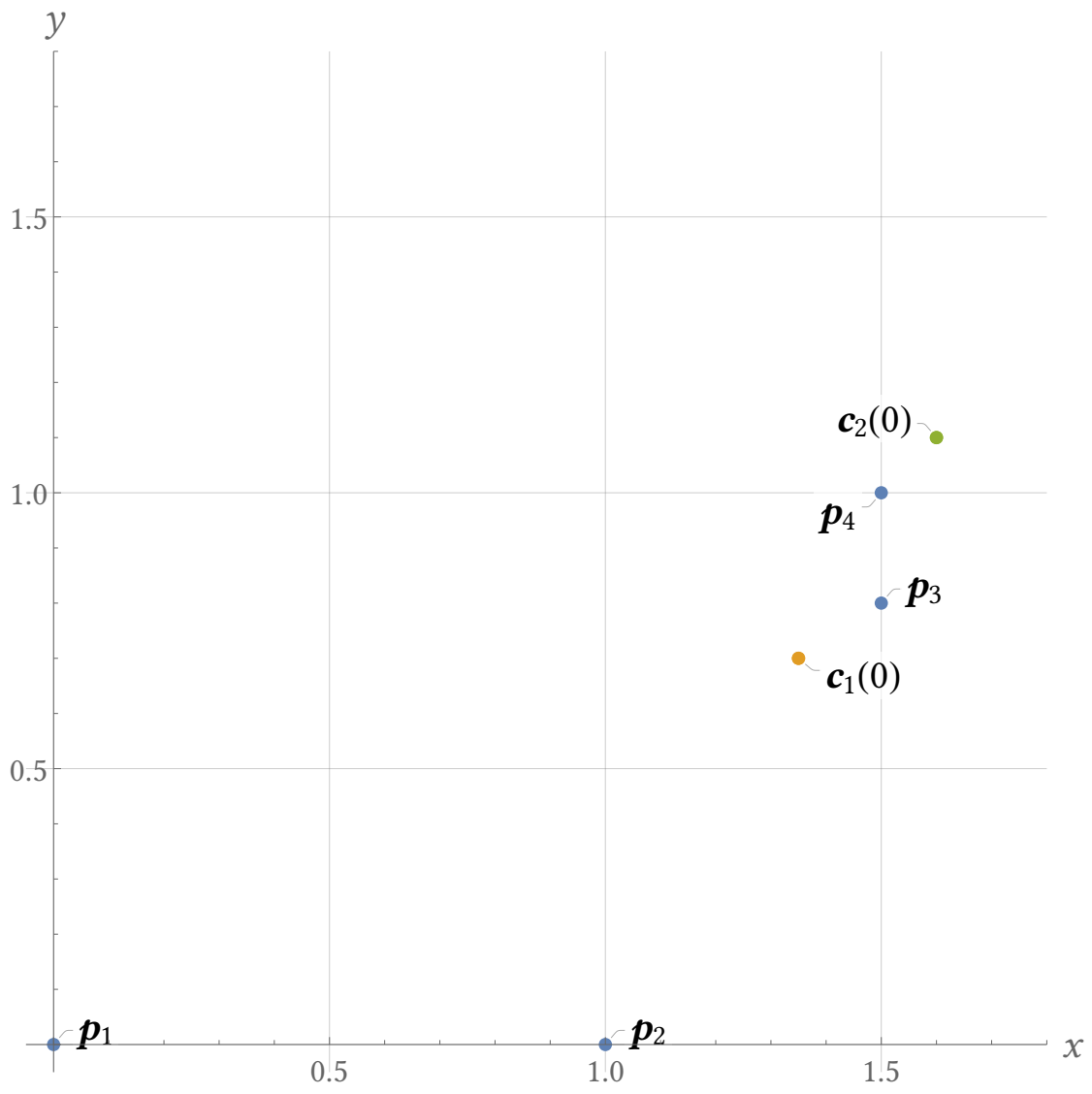
In[18]:=
```
plotBasis[t_] := ListPlot[{
   MapThread[Callout[#1, #2] &, {data, dataLabels}],
   {Callout[cluster[[1]], Row[{Subscript[bi["c"], 1], "(", t, ")"}]]},
   {Callout[cluster[[2]], Row[{Subscript[bi["c"], 2], "(", t, ")"}]]}
  },
  PlotTheme → "myTheme",
  GridLines → Automatic,
  AxesLabel → {x, y},
  AspectRatio → 1,
  PlotRange → {{-0.05, 1.8}, {-0.05, 1.8}}
 ]

plotLines[] := Show[
  Graphics[{
    ■,
    Table[
     Line[{cluster1[[i]], cluster[[1]]}]
     , {i, 1, Length[cluster1]}]
   }],
  Graphics[{
    ■,
    Table[
     Line[{cluster2[[i]], cluster[[2]]}]
     , {i, 1, Length[cluster2]}]
   }]
 ]
```

In[20]:=
```
ListPlot[MapThread[Callout[#1, #2] &, {data, dataLabels}],
 PlotTheme → "myTheme",
 GridLines → Automatic,
 AxesLabel → {x, y},
 AspectRatio → 1,
 PlotRange → {{-0.05, 1.8}, {-0.05, 1.8}}
]
```



In[21]:=
```
plotBasis[0]
```



Part 1: First Assignment

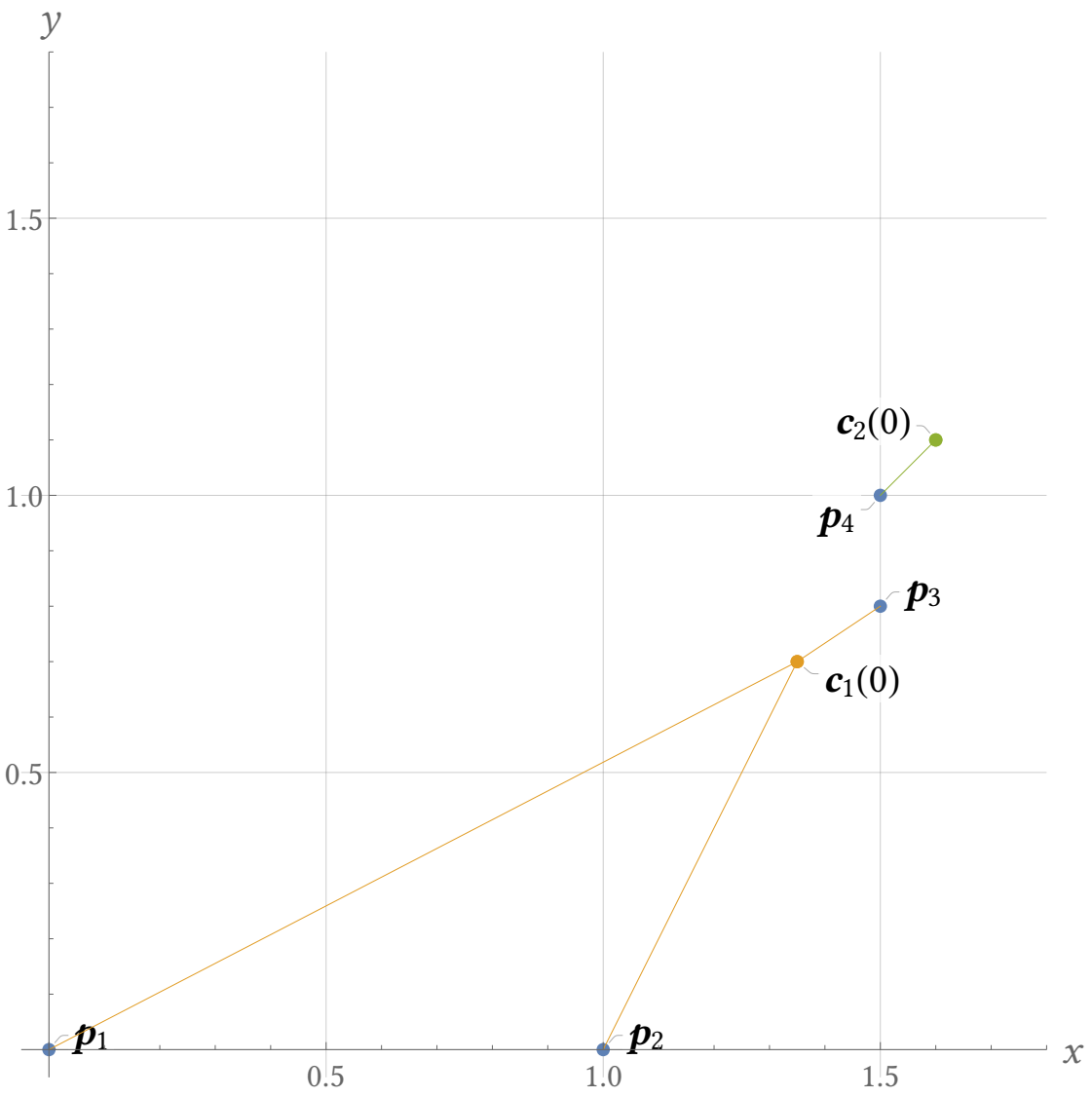
In[22]:=
```
assignPointsToClusters[]
cluster1
cluster2
```

Out[23]=
```
{{0, 0}, {1, 0}, {1.5, 0.8}}
```

Out[24]=
```
{{1.5, 1}}
```

In[25]:=
```
Show[
 plotBasis[0],
 plotLines[]
]
```

Out[25]=



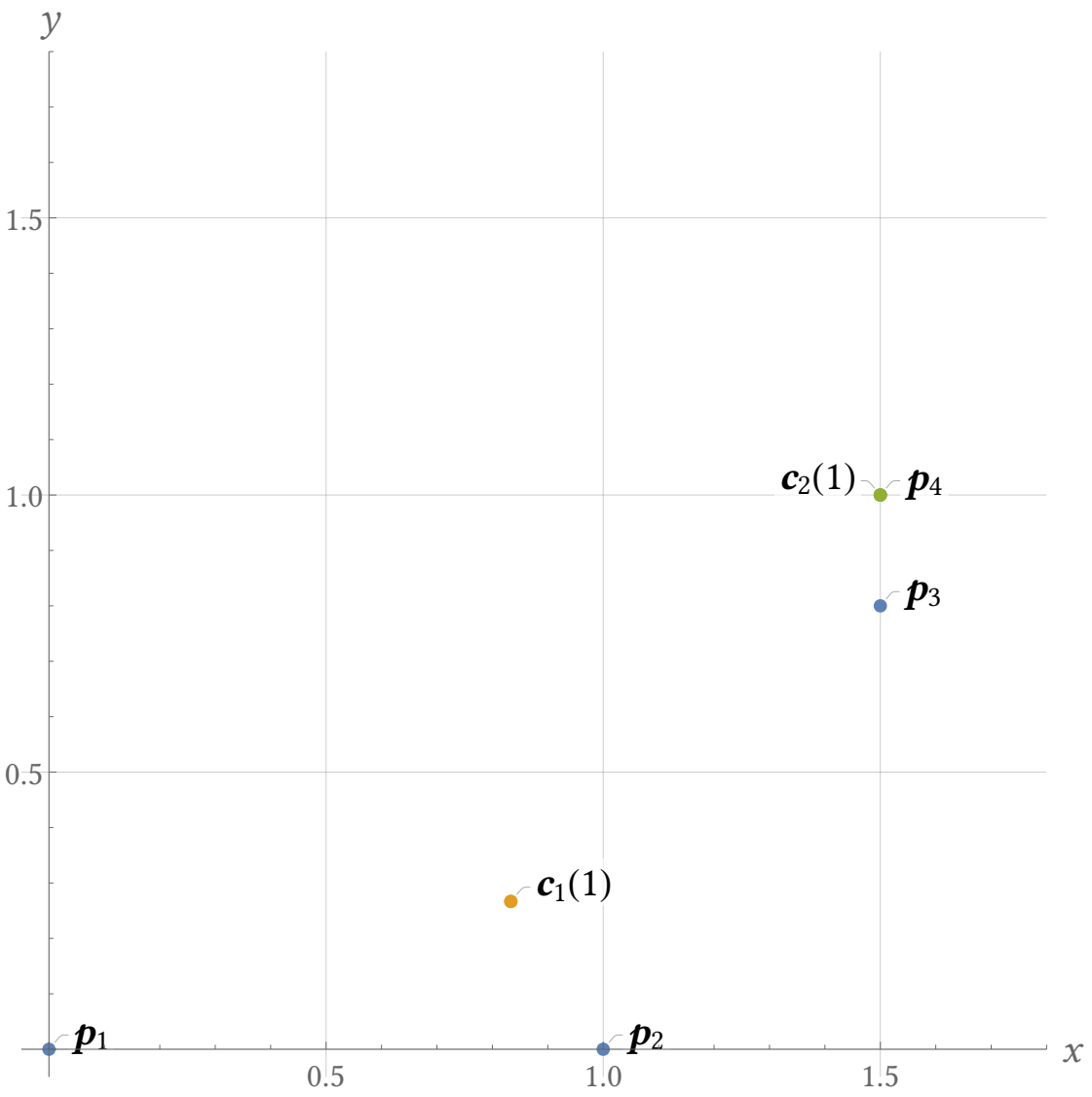
Part 2: Calculate Cluster Centre

In[26]:=
```
calculateClusters[]
cluster
```

Out[27]=
```
{{0.833333, 0.266667}, {1.5, 1}}
```

In[28]:=
```
plotBasis[1]
```

Out[28]=



Part 3: Remaining Algorithm

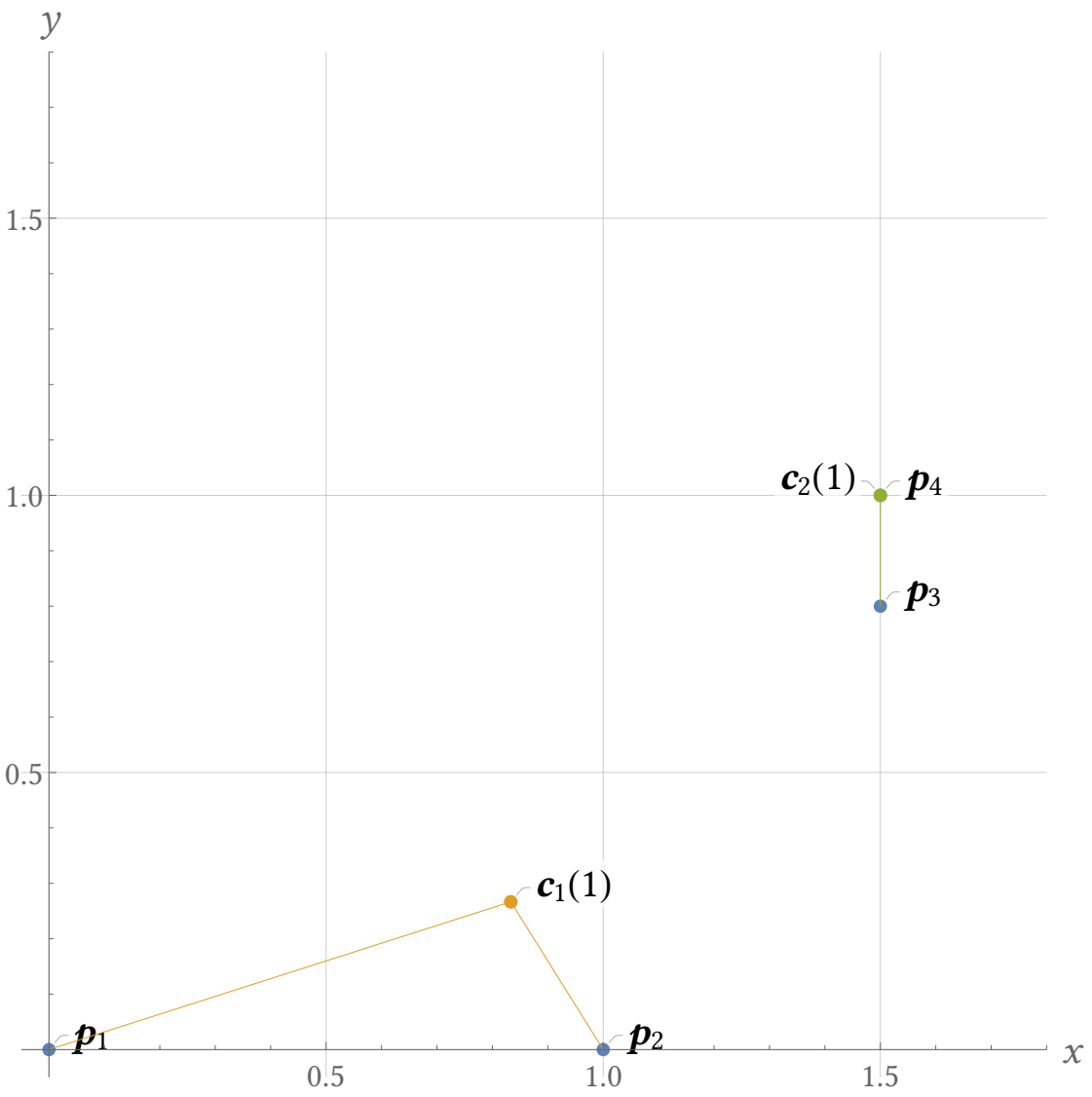
In[29]:=
```
assignPointsToClusters[]
cluster1
cluster2
```

Out[30]=
```
{{0, 0}, {1, 0}}
```

Out[31]=
```
{{1.5, 0.8}, {1.5, 1}}
```

In[32]:=
```
Show[
 plotBasis[1],
 plotLines[]
]
```

Out[32]=



In[33]:=
```
calculateClusters[]
```

In[34]:=
```
plotBasis[2]
```

Out[34]=