

参考：統計データの分析に Excel グラフを使った例

過去の課題で渋谷区や京都市の年齢別人口分布についてとりあげた。グラフ化作業を通じた簡単な分析を以下に試みたので参考にするとよい。

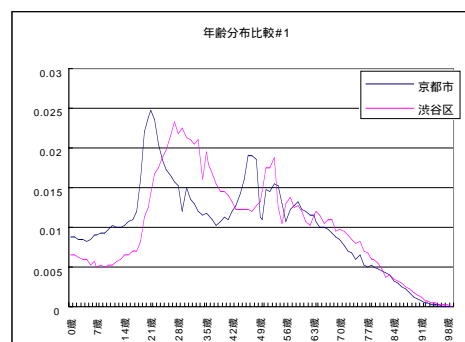
渋谷区、京都市、全国国勢調査データはそれぞれ下記から取った。

http://www.city.shibuya.tokyo.jp/ku/statics/base/jumin_age.html

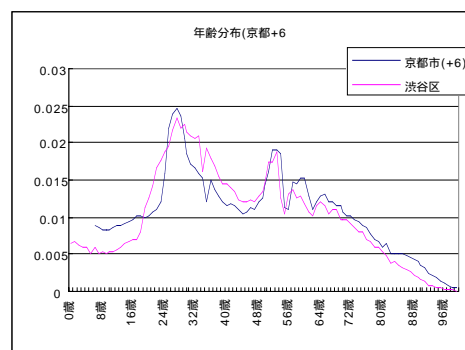
<http://www.city.kyoto.jp/sogo/toukei/census/census1995f1.html>

<http://www.stat.go.jp/data/kokusei/2000/sokuhou/zuhyou/a001-1.xls>

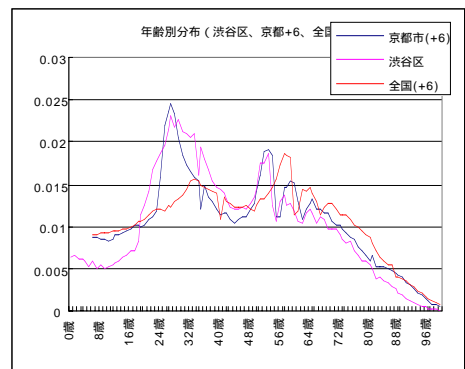
まず単純に上記データについて、渋谷区と京都市の年齢別人口分布を重ねてみた。但し、実数では京都市合計 1,454,223 渋谷区合計 190,663 と桁が違うため比較できない。そこで各年齢の全体に対するパーセンテージを算出し、比較した。(右 1 図)



奇妙に何年かのズレを保ちながら、おおよそ同じ分布をたどる。原因を調べると、そもそも元データの採集時期が異なる。前者は H7、後者は H13 である。そこで京都市人口を +6 年ずらして再描画させた。(右 2 図)



ほぼ同じカーブをたどり、20 代後半、50 代半ばにピークがある。これの原因を追及する代わりに、全国の年齢分布と比較した。データは国勢調査の全国合計から取り、同様に各年齢の全体への割合を重ねた。これも採取が H7 であるため、+6 ずらしている。(右 3 図)



これによって、青年層のピークは少なくとも全国では見られないことが判明した。このピークの原因が何か、調べる価値があることがわかる。

可視化がこのように分析の助けになることを意識すれば、Excel は更に有用なツールになることがわかるだろう。