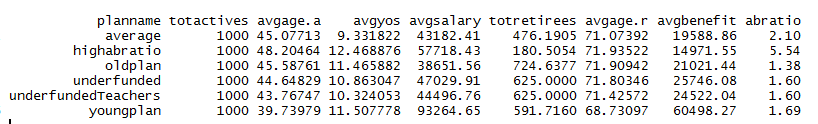
Prototype notes

7/7/2015

# Summary of protos to date



adf <- actives %>% group\_by(planname) %>%

summarise(totactives=sum(nactives),

avgage.a=sum(age\*nactives) / totactives,

avgyos=sum((age-ea)\*nactives) / totactives,

avgsalary=sum(salary\*nactives)/totactives)

rdf <- retirees %>% group\_by(planname) %>%

summarise(totretirees=sum(nretirees),

avgage.r=sum(age\*nretirees) / totretirees,

avgbenefit=sum(benefit\*nretirees) / totretirees)

left\_join(adf, rdf) %>% mutate(abratio=totactives / totretirees)

# Young and old plans

## Conclusions

* Base very high abratio plan on WA PERS 2 (ppd 119)
* Base very old plan on MI public school (ppd 53)
* Base very young plan on NJPFRS (ppd 72)

## Investigation

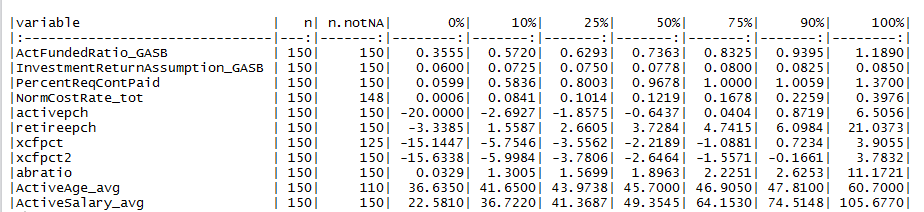
5 variables of interest:

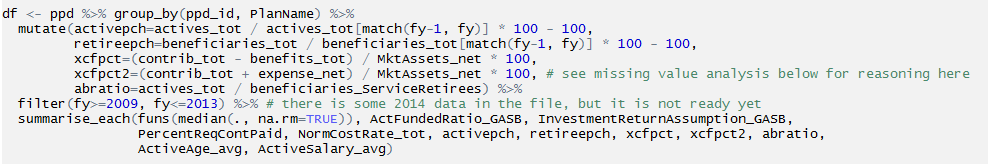
* Average age of actives (has about 90 obs recent years)
* Average age of retirees (too many missing obs – only about 40 obs)
* Growth rate number of actives
* Growth rate number of retirees (use beneficiaries – slightly more data); in retrospect (after I am well into this), retiree growth as % of total members might be a better measure – an old plan could have rapid retiree growth as many older workers move into retirement, but a younger plan also might have rapid retiree growth if it starts out with very few retirees; but in this case, retiree growth as % of total members would be low while in former case it would be high
* Ratio of actives to retirees (abratio – actives to beneficiaries)

Possible partial proxy for age of plan could be external cash flow as % of assets (xcfpct2) but of course it is related to funded ratio and contribution percent.

### Construct variables and get basic stats

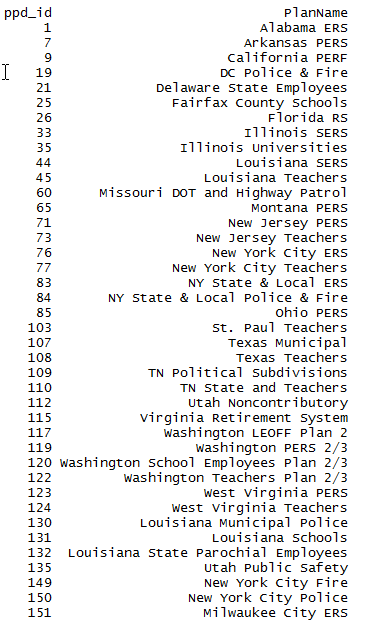
Construct growths and ratio; get 2009-2013 medians for all, to get more data





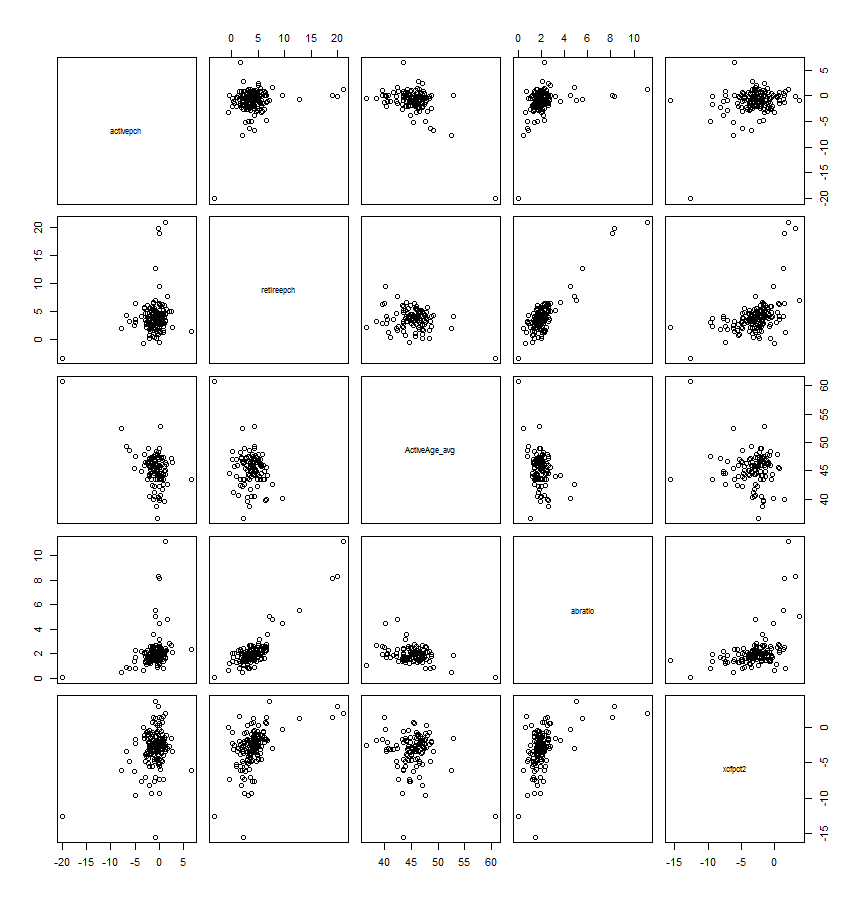
### Maybe we’ll have enough obs for active age?

Here are the 40 plans missing active age in the constructed (medians) data set. Missing CalPERS, NY plans, NJ plans, some IL plans, so probably we’ll have to exclude actives age



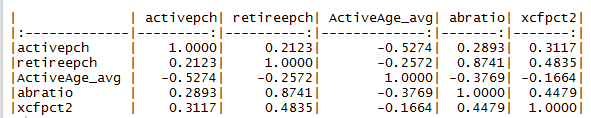
### Scatters

Retireepch and abratio; maybe a little with xcfpct2



### Correlations

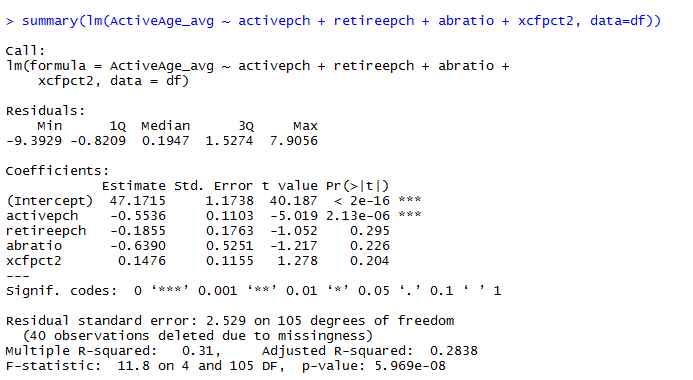
Pairwise complete obs



A little surprised the correlation on active-retiree pch pairing is positive

Active age and active pch fairly highly correlated

Is active age conditionally correlated with other variables?

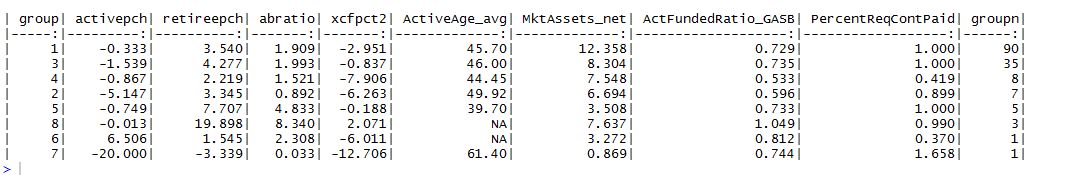


### Cluster analysis

Leave active age out of the clustering, but do look at it

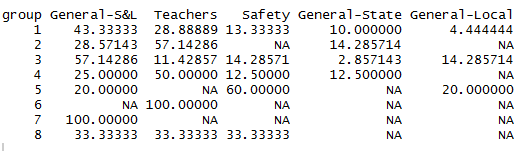
Summary of approach

* I used 4 of the 5 variables for clustering – not actives age because too many missing values
* Scaled variables to 0 mean, 1 SD
* Calculated Euclidean distance of each to each
* Did hierarchical clustering with defaults
* Selected 8 clusters (no magic to this – enough to allow clear groups and to let some of the outlier categories have their own groups)
* Merged in selected non cluster variables from the ppd and calculated median of 4 cluster variables plus selected non cluster variables for each group (these variables for given plans already were the plan medians over 2009-2013, so this is the median of each of those variables)
* Here are the results sorted by # of plans in group, followed by medians for all 150 plans:





* Group 1 is a “typical” plan group and has the most plans; disproportionately general employees (table below)
* Group 3 also looks pretty typical plan group but slightly older and getting older faster (activepch more negative, retireepch more positive) moderately older plan; disproportionately general; AZ-SERS (ppd #6), our average plan, is in here
* Group 4 is a younger slow-growing plan group (but abratio is low -- ??); it is dominated by teacher plans (see below) but it only has 8 plans
* Group 2 is a very old plan group; rapid actives declines, rapid retiree growth, old average age, low abratio; it, too, is dominated by teacher plans but there are only 7 plans in it; CA PERF is in here
* Group 5 is a very young plan group (but we only have age for one plan in the group), slow declines in actives (? Could they be closed plans – check before using as a model?), but it has fast retiree growth (maybe that’s ok if starting from a low base), high abratio
* Here are the relative frequencies by plan type

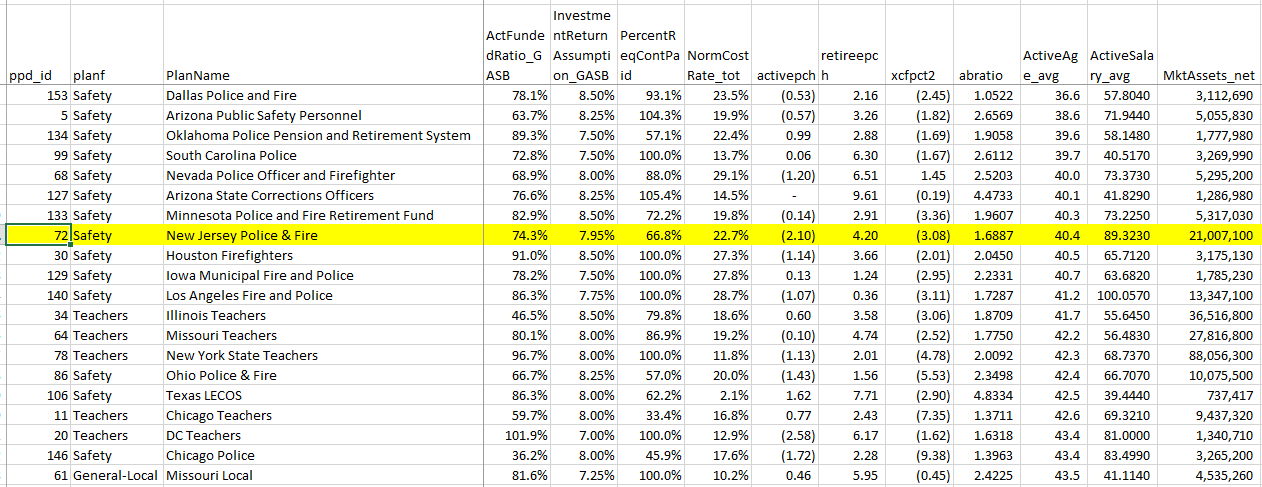


## Selecting a plan a basis for young prototype and old prototype

Plans in selected groups, sorted by assets

### Younger plans – brute force

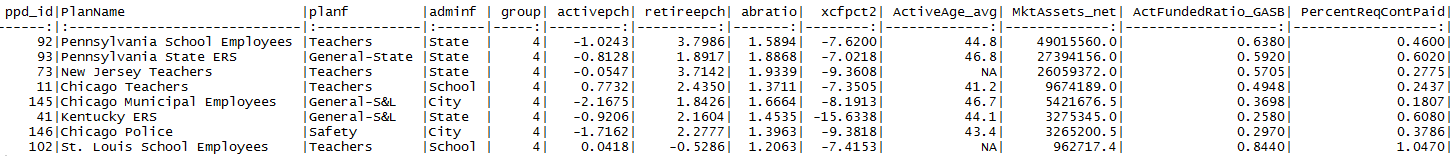
Here are the 20 youngest plans. I picked NJ police and fire. It is young, but not immature.



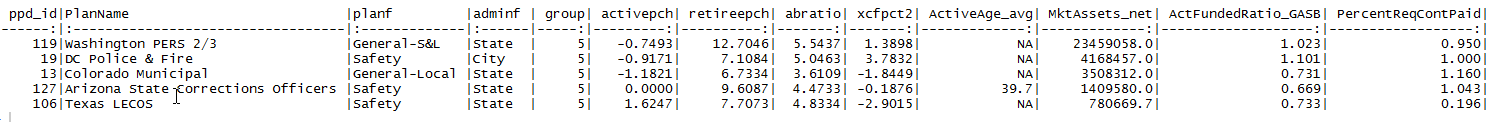
### Younger plans – this ended up with a high abratio plan

#### Group 4 younger

Note that PA-PSERS, our poorly funded high outflows prototype is in this group. I don’t really think of these plans as young plans. And they have relatively low abratios I think.



#### Group 5 very young



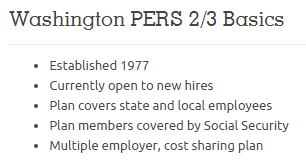
#### Candidates

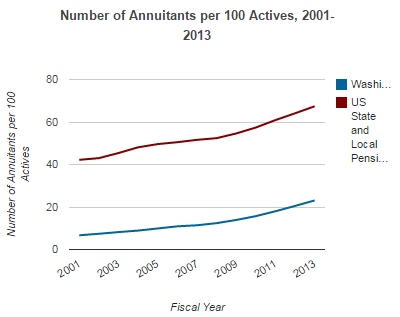
WA PERS2 looks like a good candidate to me, but we don’t know avg age. Or possibly AZ state corrections.

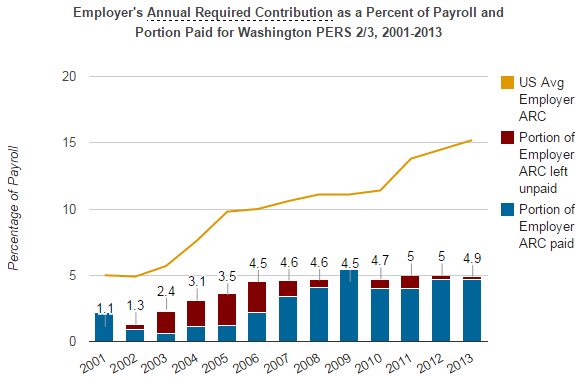
#### Washington PERS as possible young plan prototype; selected info

<http://publicplansdata.org/quick-facts/by-pension-plan/plan/?ppd_id=119>

open; well-funded; low ER NC; high ERC/ARC; positive xcf;



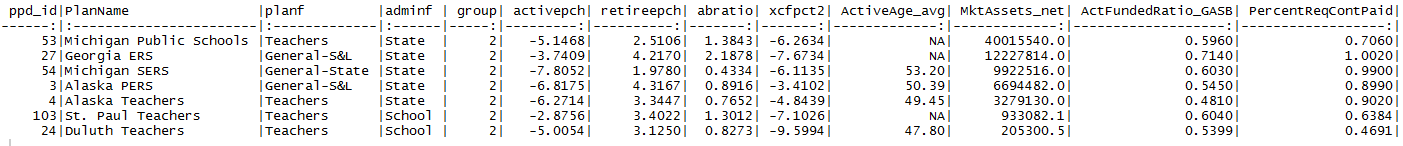




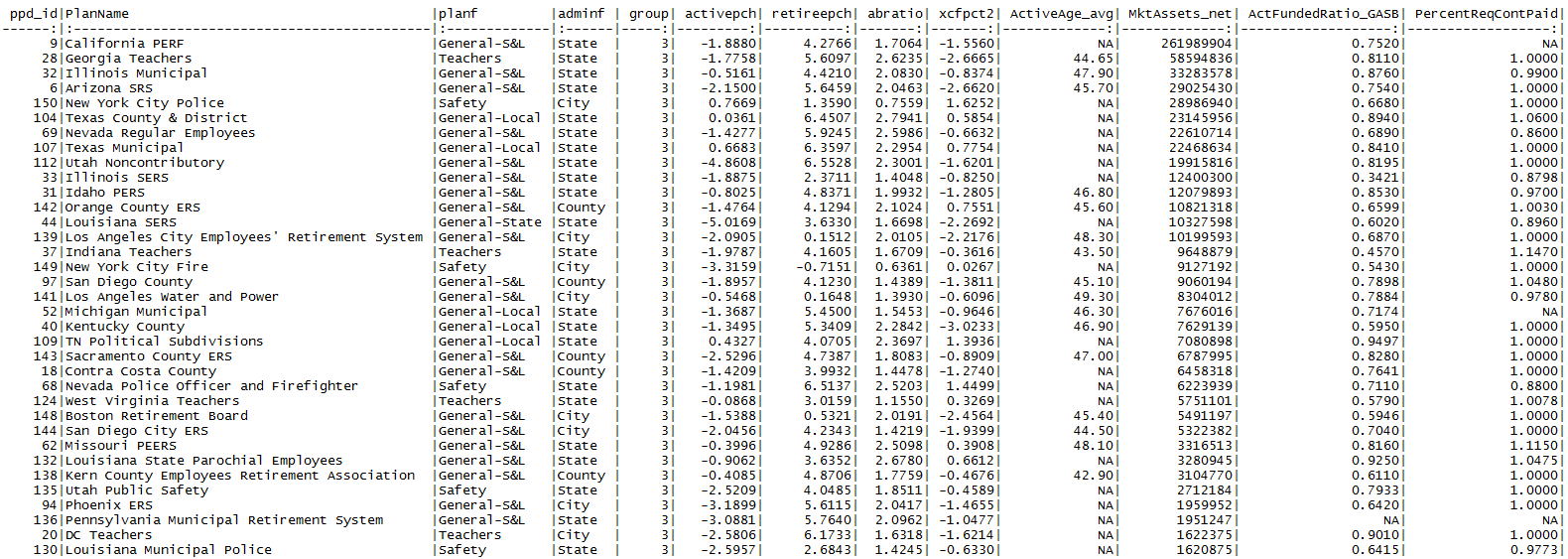
Looks ok to me

### Older plans

#### Group 2 very old



#### Group 3 slightly old



#### Candidates

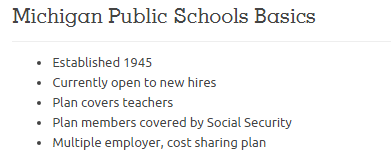
On discussion with Yimeng, MI public schools looks like a good plan

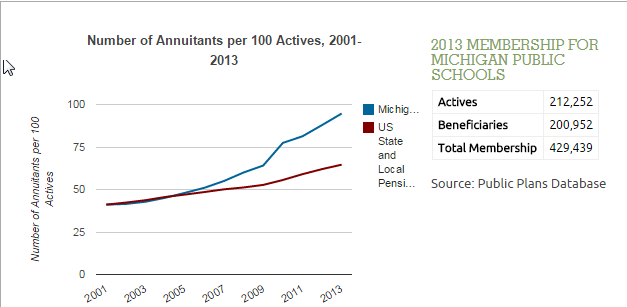
St. Paul Teachers from group 2 looks like a good candidate to me. Or perhaps Duluth, but it is much smaller.

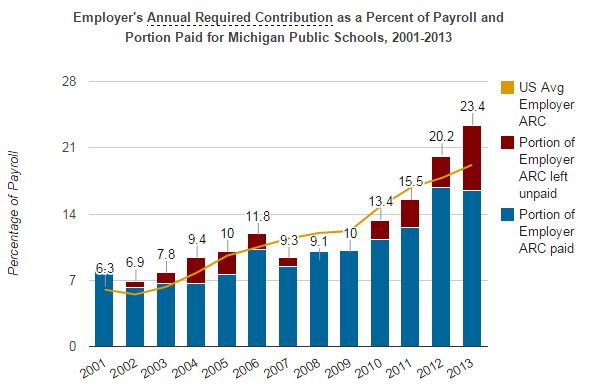
Cal PERF from group 3 might also be a good candidate, perhaps aged based on recent trends.

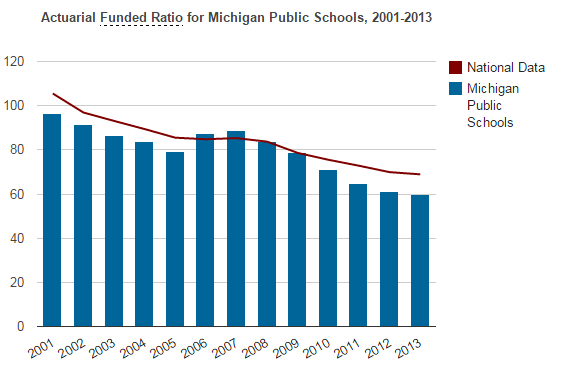
#### MI School as a candidate

A lot of arc underfunding, but our main concern here is demographics.









#### St. Paul Teachers as a candidate

Poor ARC payment history, low funded status

