

## Notes for Creating Complex Files with PDQ Explore

### Introduction

PDQ Explore allows one to pull off extracts from microdata files very quickly. However, the more valuable advantage is that it allows one to create complex files during the extraction process. Examples of complex files would be files where two different records are combined into a single record such as husband/wife, head/partner, and mother/child.

These notes assume the user is familiar with PDQ Explore and has used it to create an extract.

Selected complex custom items are defined within PDQ. See us for help if you have a complex custom item in mind.

Defined in PDQ are:

Husband/Wife  
*PUMS*      *1990 and 2000*  
*March CPS*    *1990–2003*

Head/Partner [unmarried partners]  
*PUMS*      *2000*

Child/Father/Mother  
*PUMS*      *1990 and 2000*

The following are some simple examples. The examples show the universe description, item selections, and what the extract will look like.

### **Husband/Wife extract [2000 PUMS]**

#### ***Universe:***

`msp=1 & sex=1`

#### ***Item selections:***

```
state
serialno
pnum
relate
sex
msp
age
educ
earns
wife(serialno)
wife(pnum)
wife(relate)
wife(sex)
wife(msp)
```

```
wife(age)
wife(educ)
wife(earns)
```

**Sample Results:**

```
26 00001 1 1 1 31 10 50100 00002 2 2 1 30 11 38765
26 00002 2 1 1 22 9 22456 00001 1 2 1 20 9 30000
```

**Notes**

The user needs to define the universe of one of the persons in the pair. In a husband/wife file the universe should be limited to married spouse present and either male or female. If the user uses the “husband” custom item, then the universe should be defined for the wife, e.g., married spouse present and female. If the user uses the “wife” custom item, the universe should be defined for the husband – married spouse present and male.

Once the universe is defined, the items are selected for the “universe defined person” and then for the partner.

In the example above, the universe is defined for the husband and the “wife” custom item is used.

**Child and Parents extract [2000 PUMS]**

**Universe:**

```
age>=8 & age<=10
```

**Item selections:**

```
state
sum( (age=8) : 0..29 )
sum( (age=9) : 0..29 )
sum( (age=10) : 0..29 )
relate
age
sex
racel
father(age)
father(racel)
father(educ)
father(earns)
mother(age)
mother(racel)
mother(educ)
mother(earns)
```

**Sample Results:**

```
26 1 0 0 3 8 1 1 31 1 10 47500 30 1 10 40500
26 0 1 0 3 9 2 2 -1 -1 -1 -1 32 2 10 31700
26 0 1 0 8 9 1 0 0 0 0 0 0 0 0 0
```

## Notes

The first child has a father and a mother. The household information is in items 1-4, the child information is in items 5-8, the father information is in items 9-12, and the mother information is in items 13-16. The second child has a mother but no father. The fields that represent the father have missing data codes. The third child has no information about her parents on her record because she is not living with any of her parents. She lives with her grandparents.

The user needs to examine the raw output to determine how to assign missing value codes with a statistical package.

The 'sum' items on the item selections are creating a count of how many 8 year olds, 9 year olds, and 10 year olds are in the household. This is useful to know for weighting purposes in case a household is represented more than once in the extract, e.g., there are twins or two closely spaced siblings.

## Unmarried partner [PUMS 2000]

### *Universe:*

```
relate=1 & msp>2 & up(age)>0
```

### *Item selection:*

```
sex  
age  
up(sex)  
up(age)
```

### *Sample Results:*

```
1 23 2 21  
1 40 2 38  
2 26 1 30  
1 24 1 22  
2 57 2 55
```

## Note

The census makes no assumptions about the gender of the head of the household. If one is studying heterosexual unmarried partners, one will have to re-arrange the extract so that the male partner is in the same location in the file across all records. One can do this by running two separate extracts in PDQ Explore (shown below).

The extract above includes all unmarried partners (same sex and heterosexual). In the statistical package one would re-arrange the order of the partners, if desired.

The up(age)>0 universe selection is included to make sure that only households that have an unmarried partner are included in the output.

**Universe:**

relate=1 & sex=1 & msp>2 & up(sex)=2

**Item selection:**

sex  
age  
up(sex)  
up(age)

**Sample Results:**

1 23 2 21  
1 40 2 38  
1 19 2 19  
1 44 2 42  
1 51 2 53

**Universe:**

relate=1 & sex=2 & msp>2 & up(sex)=1

**Item selection:**

up(sex)  
up(age)  
sex  
age

**Sample Results:**

1 30 2 26  
1 51 2 49