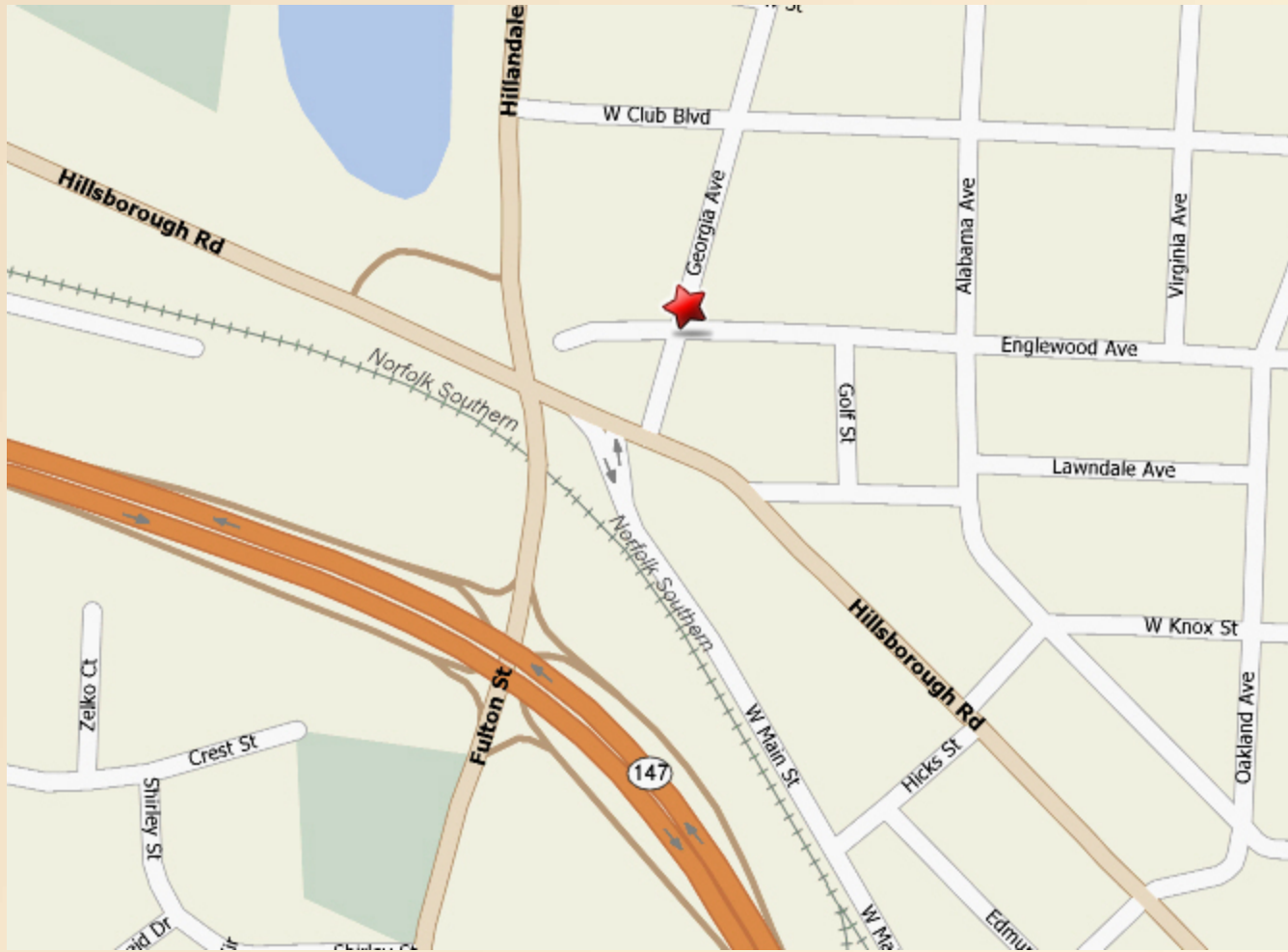




New in SAS 9.2: It's the Little Things that Count

Diane Olson
Senior Systems Developer, SAS Institute Inc.



Data Management

Cross-host Data Migration

- CEDA (Cross-Environment Data Access) is key
- Catalog access via Remote Library Services

PROC MIGRATE Syntax

```
libname lib1 v8 'physical location 1';  
libname lib2 base 'physical location 2';  
libname cats server=server1;  
proc migrate in=lib1 out=lib2  
  slibref=cats;run;
```

REBUILD an Index

- `OPTION DLDMGACTION=NOINDEX;`
- Indexes and integrity constraints marked inactive
- Repaired data set is read-only until REBUILD is used

REBUILD Statement

```
proc datasets data= SAS-library;  
  rebuild SAS-file <noindex>; run;  
quit;
```

IBUFNO System Option

- Sets number of extra buffers used by indexes
- Increase index performance vs. increased memory usage
- Options IBUFNO= < 0 – 10,000 >;
- Use in conjunction with IBUFSIZE

Data Creation

SET Data sets: The Hard Way

data results;

set a1 a2 a3 a4 a5 a6 a7 a8 a9 a10

a11 a12 a13 a14 a15 a16 a17 a18 a19 a20

a21 a22 a23 a24 a25 a26 a27 a28 a29 a30

a31 a32 a33 a34 a35 a36 a37 a38 a39 a40

a41 a42 a43 a44 a45 a46 a47 a48 a49 a50

a51 a52 a53 a54 a55 a56 a57 a58 a59 a60

a61 a62 a63 a64 a65 a66 a67 a68 a69 a70

a71 a72 a73 a74 a75 a76 a77 a78 a79 a80

a81 a82 a83 a84 a85 a86 a87 a88 a89 a90

a91 a92 a93 a94 a95 a96 a97 a98 a99 a100;

run;

Data Set Lists

- Dash list syntax

A1-A100 A1B1 –A1B5 A100-A1

Data Set Lists

- Dash list syntax

A1-A100 A1B1 –A1B5 A100-A1

- Colon list syntax

A: JUNE: CLASS:

SET Data Sets: Easy Ways

```
data results;  
    set a1-a100;  
run;
```

```
data results;  
    set a::;  
run;
```

STEP Statement with INDSNAME Option

```
data results;
```

```
    set gas_price_option
```

```
        gas_rbid_option
```

```
        coal_price_option
```

```
        coal_rbid_option
```

```
    indsname = cur_dataset;
```

```
    from_data_set = cur_dataset ;run;
```

```
proc print data=results; run;
```

Proc Print with INDSNAME Results

Obs	x	from_data_set
1	1	WORK.GAS_PRICE_OPTION
2	2	WORK.GAS_RBID_OPTION
3	3	WORK.GAS_RBID_FORWARD
4	4	WORK.COAL_PRICE_OPTION
5	5	WORK.COAL_RBID_OPTION
6	6	WORK.COAL_RBID_FORWARD

INDSNAME Option – Parse Data Set Name

```
data results;
```

```
  set gas_price_option ...
```

```
  indsname = cur_dataset;
```

```
  fuel = scan(cur_dataset, 2, '._');
```

```
  value = scan(cur_dataset, 3, '._');
```

```
  type = scan(cur_dataset, 4, '._'); run;
```

Proc Print Results of Parsing Data Set Name

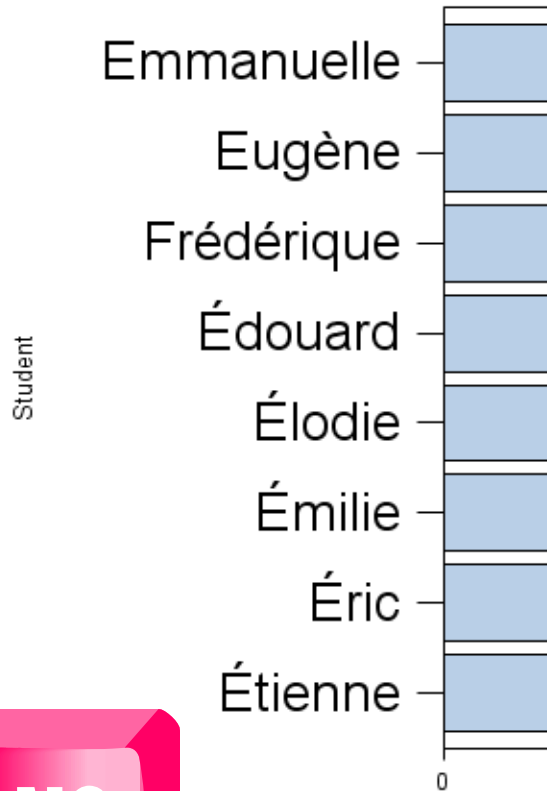
Obs	x	fuel	value	type
1	1	GAS	PRICE	OPTION
2	2	GAS	RBID	OPTION
3	3	GAS	RBID	FORWARD
4	4	COAL	PRICE	OPTION
5	5	COAL	RBID	OPTION
6	6	COAL	RBID	FORWARD

Data Organization

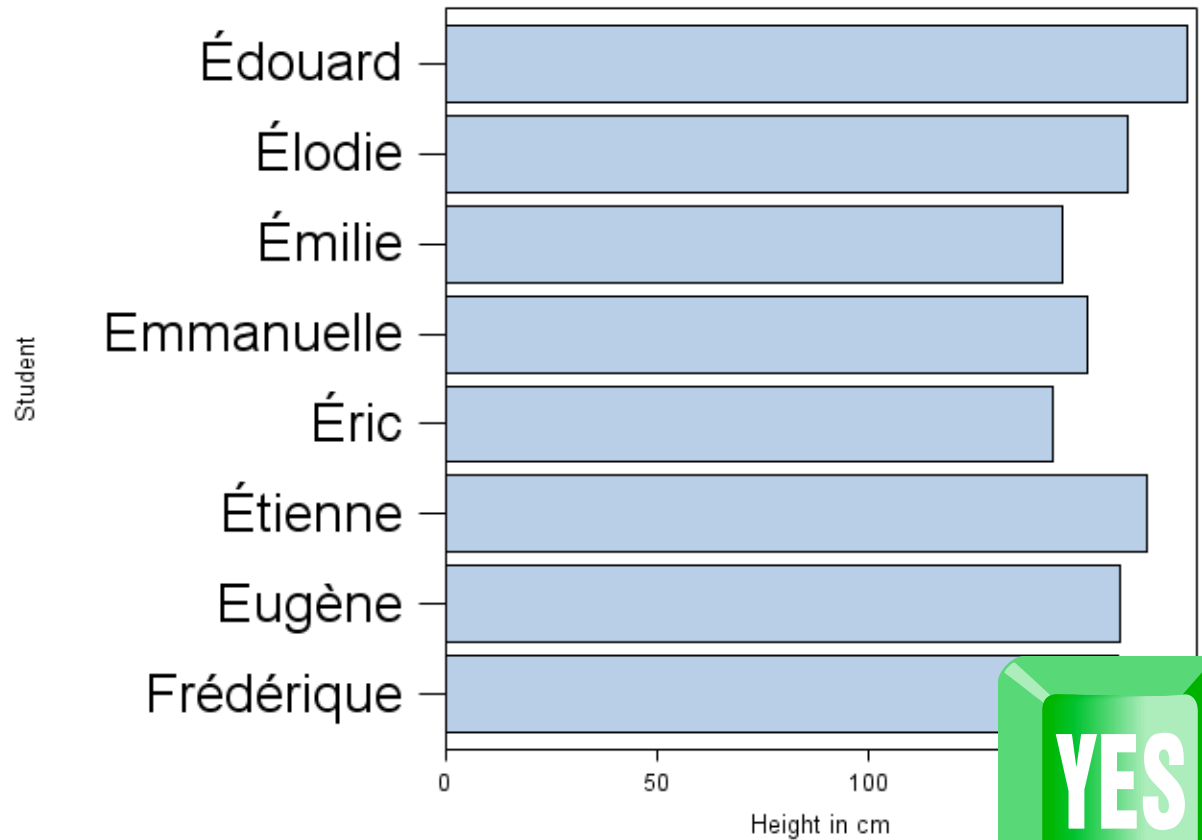
Culturally Correct Sorting

- Outside of United States
- Inside United States
- Numeric Collation

Student Height



Student Height



Linguistic Sorting

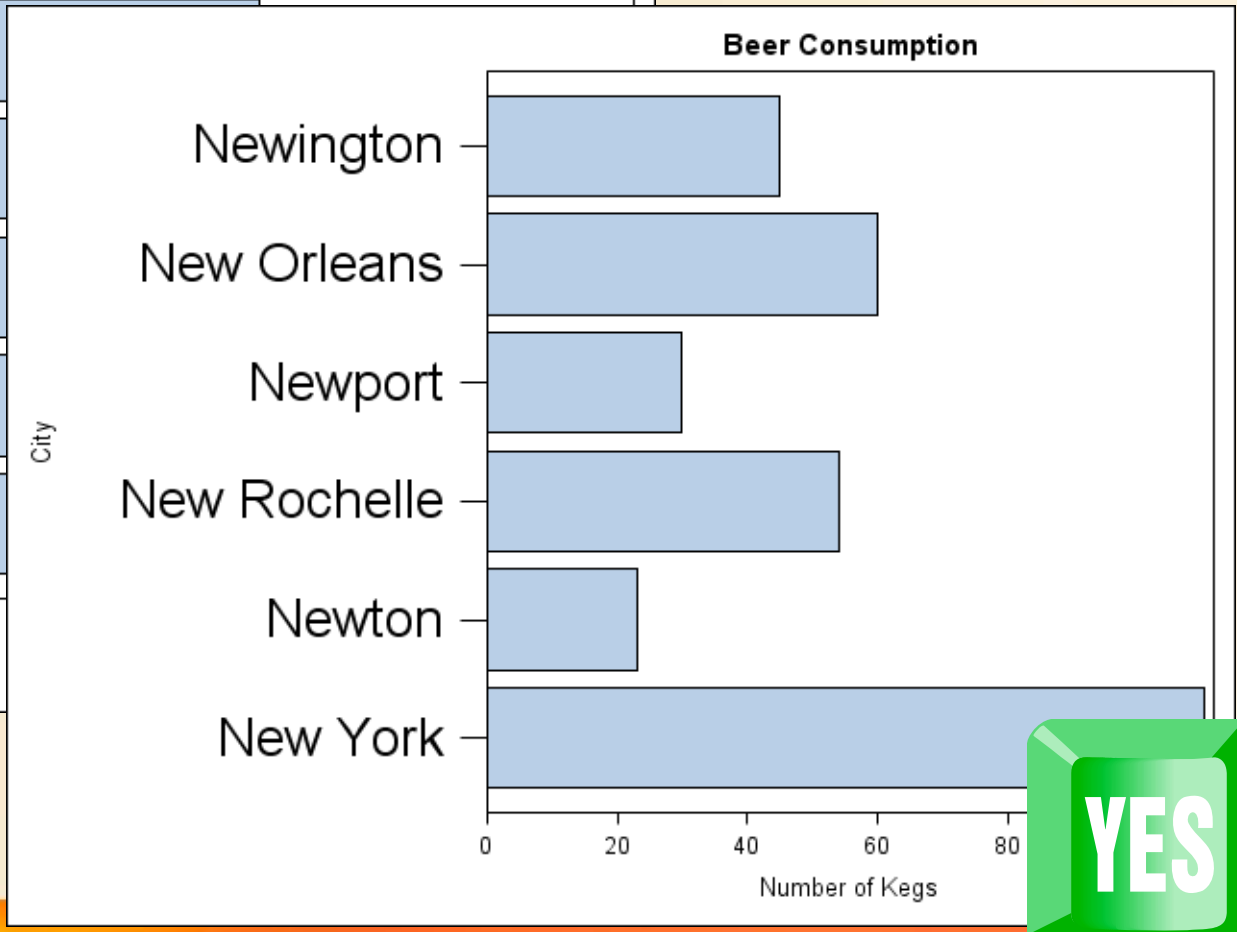
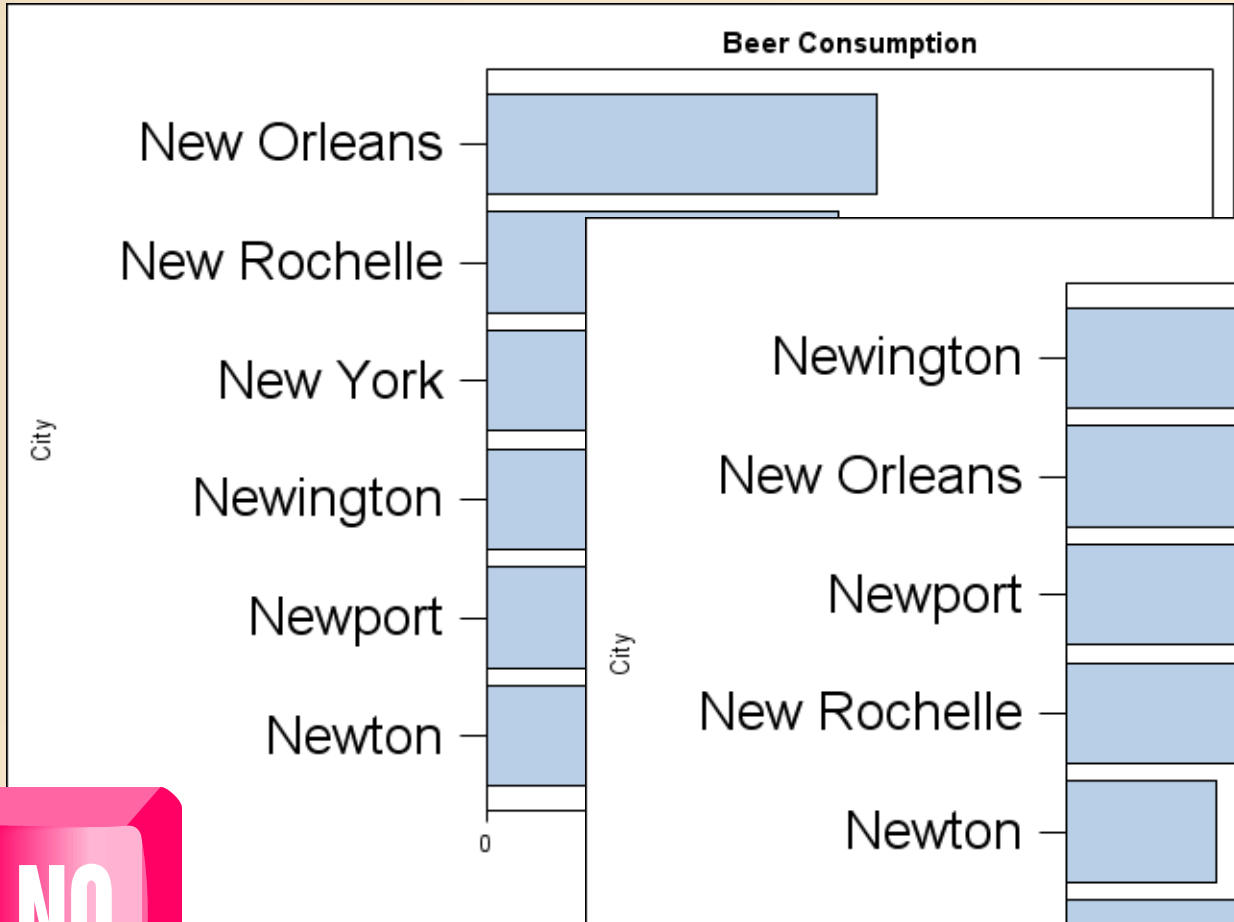
```
proc sort data=class
      out=sorted
      sortseq=
          linguistic(locale=fr_FR) ;
      by name ;
run ;
```

Linguistic Sorting

Obs	Binary	Lowfirst	Linguistic
1	Aaron	aardvark	aardvark
2	Aztec	azimuth	Aaron
3	Zeus	Aaron	azimuth
4	aardvark	Aztec	Aztec
5	azimuth	zebra	zebra
6	zebra	Zeus	Zeus

Linguistic Sorting

```
proc sort data=customers  
          sortseq=linguistic;  
  by name;  
run;
```



Spaces and Symbols Less Important

```
proc sort data=consumption
out=sorted
sortseq=
linguistic(
alternate_handling=shifted);

by city;

run;
```

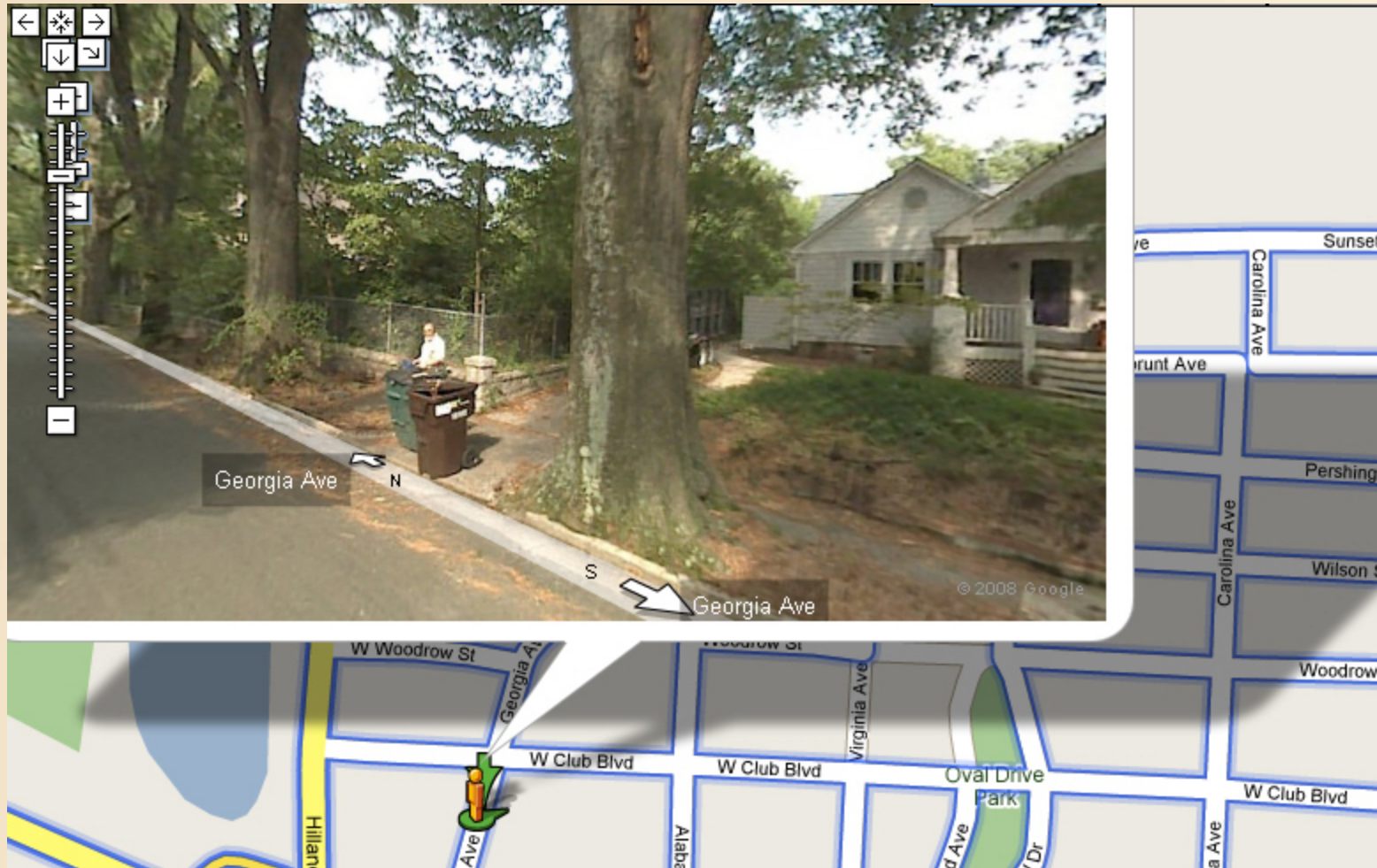
Numeric_Collation=ON

Obs	Binary	Linguistic
1	0123 Main St. Apt #12	123 Main St. Apt #1
2	123 Main St. Apt #1	123 Main St. Apt #2
3	123 Main St. Apt #103	0123 Main St. Apt #12
4	123 Main St. Apt #2	123 Main St. Apt #24
5	123 Main St. Apt #24	123 Main St. Apt #103

Proc Contents – Sort Information

Sortedby	X
Validated	YES
Character Set	ANSI
Collating Sequence	LINGUISTIC
Locale	en_US
Strength	3





Questions?

Diane.Olson@sas.com



2008 SAN ANTONIO TEXAS

16-19 MARCH 2008