

Hwk #3: Data Visualization & More

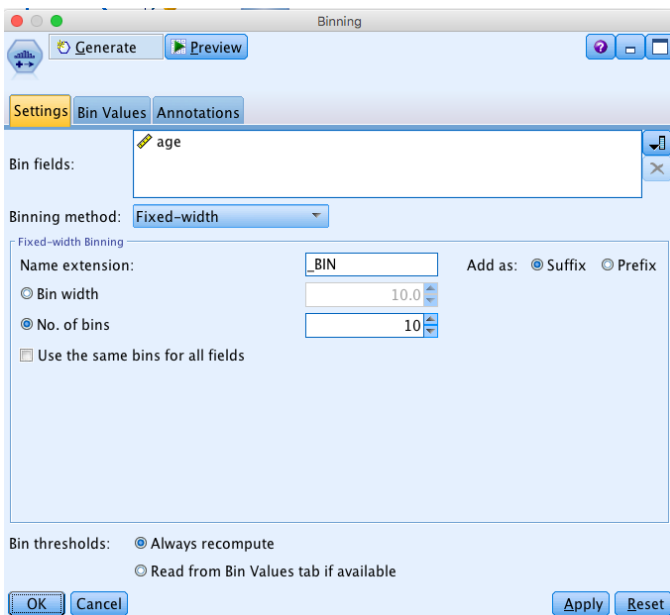
Biostat 202: Opportunities and Challenges of “Big Data”

Due Wednesday, August 16th by Midnight

For this homework we will also use the data from the National Electronic Injury Surveillance System (NEISS) that you used for Hwk #1 (<http://www.cpsc.gov/en/Research--Statistics/NEISS-Injury-Data/>). We will be interested in creating plots of the data and examining how to evaluate models graphically.

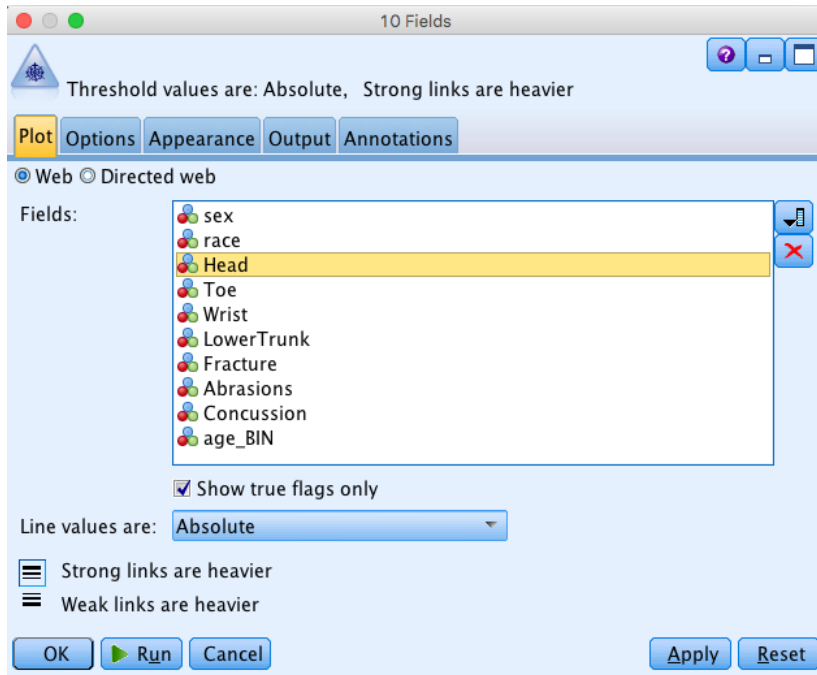
Open your stream from Hwk #2. I have also included the data as an Excel file in Hwk #2 if you prefer to use this instead of the .csv file. You should have the recoded variables for admit, Head, Toe and Age already in your Type node.

1. (3 points) Using the codes in the table below from the NEISS codebook, choose two more ‘Body Part Affected’ codes and create new variables as you did in Hwk #2 for Head & Toe. **Q1. Which areas of the body did you choose? Do you think that injuries to these new areas are likely to involve hospitalization?**
2. (3 points) Using the codes in the table below from the NEISS codebook, choose three diagnoses from the ‘Diagnosis’ table and create variables. **Q2. Which three diagnoses did you choose? Do you think any of these are related to the body parts chosen above?**
3. (5 points) Attach a Binning node from Field Ops and create Age groups to use in analyses that will not accept continuous data. Create 10 bins (or groups)

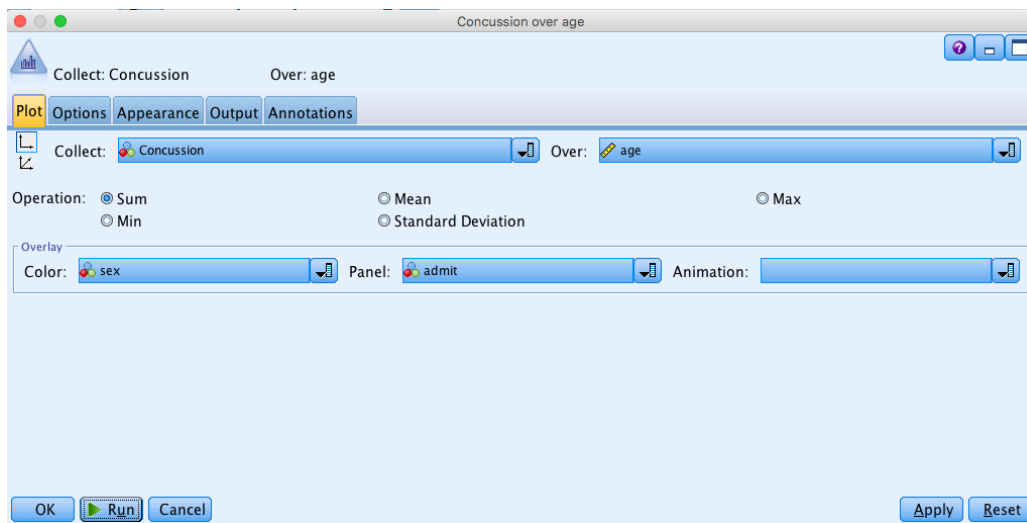


Attach a Type node following all your reclassification nodes and binning node. For the plots and models, we will use the following variables: a continuous Age field ranging from 1 to 107, an admit variable (0/1), 4 body part variables (0/1), 3 diagnosis variables (0/1), a gender variable (0, 1, 2), and an Age_bin for our age groups. Run an Audit node and then get a distribution of your new age_BIN variable by double clicking on the bar chart for the variable. **Q3. By binning by number of bins rather than the actual distribution of Age, what is this distribution telling you about the population?**

4. (4 points) Attach a web graph to your Type node and include the variables above (except Age). Use the bar below the web to examine how the variables are related. **Q4. Which variables show relationships? Are they variables that occur together or that do not occur together? Why would links between diagnoses and body parts that are not related be stronger than those that are related?**



5. (5 points) Create a Collection graph (from the Graph tab): Collect Concussion Over Age and Overlay Sex (Color) and Panel admit. **Q5. What does this tell you about concussions by age and admission to the hospital?**



6. (4 points) Finally, fit a Logistic model and a Neural net with admit as the Target and using your new diagnosis and body_part variables as well as Age and Sex. Attach Analysis nodes and Evaluation graphs to each model. **Q6. Are there differences in the accuracy of the models? Do the Evaluation plots differ? Are these models better than those fit for last week's homework? Attach the Neural model gem to the Logistic model gem and run an Analysis node. This will tell you how much your 2 models agree. Do they agree? You can also do this with your Evaluation plots.**

For #1

<u>Body Part Affected</u>	<u>Code</u>
Arm, lower (<u>not</u> including elbow or wrist)	33
Arm, upper	80
Ankle	37
Ear	94
Elbow	32
Eyeball	77
Face (including eyelid, eye area and nose)	76
Finger	92
Foot	83
Hand	82
Head	75
Internal (use with aspiration and ingestion)	00
Knee	35
Leg, lower (<u>not</u> including knee or ankle)	36
Leg, upper	81
Mouth (including lips, tongue and teeth)	88
Neck	89
Pubic region	38
Shoulder (including clavicle, collarbone)	30
Toe	93
Trunk, lower	79
Trunk, upper (<u>not</u> including shoulders)	31
Wrist	34
25-50% of body	84
All parts of body (more than 50% of body)	85
Not recorded	87

<u>Diagnosis</u>	<u>Diagnosis Code</u>	<u>Body Part Code</u>
Anoxia	65	85
Electric shock	67	85
Poisoning	68	85
Submersion (incl. drowning)	69	85

For #2

<u>Diagnosis</u>	<u>Code</u>
Amputation	50
Anoxia	65
Aspirated foreign object	42
Avulsion	72
Burns, scald (from hot liquids or steam)	48
Burns, thermal (from flames or hot surface)	51
Burns, chemical (caustics, etc.)	49
Burns, radiation (includes all cell damage by ultraviolet, x-rays, microwaves, laser beam, radioactive materials, etc.)	73
Burns, electrical	46
Burns, not specified	47
Concussions	52
Contusions, Abrasions	53
Crushing	54
Dental injury	60
Dermatitis, Conjunctivitis	74
Dislocation	55
Electric shock	67
Foreign body	56
Fracture	57
Hematoma	58
Hemorrhage	66
Ingested foreign object	41
Internal organ injury	62
Laceration	59
Nerve damage	61
Poisoning	68
Puncture	63
Strain or Sprain	64
Submersion (including Drowning)	69
Other/Not Stated	71