

# Graph Theoretic General Linear Model (GTG): a MATLAB toolbox

# Jeffrey M. Spielberg VA Boston Healthcare System & Boston University School of Medicine

## **Graph Theory & Psychopathology**

There is increasing agreement that that psychopathology is best understood as *disturbances in brain networks* (Insel, 2010). The recent application of graph theory to clinical neuroscience allows network pathology to be studied in an increasingly sophisticated manner. Graph-theoretic topological properties provide insight into the organization of networks and the role of regions within a network, and can reduce the incredibly large search space of brain networks in a meaningful manner (Rubinov & Sporns, 2010). Thus, graph theory has the potential to greatly expand knowledge of pathology-related network disturbance, enriching our understanding of etiological pathways.

Categories of topological graph properties include:

**Segregation** – the degree of optimization for specialized processing to occur within densely interconnected groups of nodes. Properties include: *Clustering Coefficient, Local Efficiency, Transitivity* 

**Integration** – the degree to which a network can rapidly combine specialized information across distributed nodes. Properties include: Characteristic Path Length, Global Efficiency

*Centrality (Influence)* – the degree to which a particular node/edge facilitates intercommunication. Properties include: Node Degree/ Strength, Node/Edge Betweeness, Eigenvector Centrality, Pagerank Centrality, K-Coreness Centrality, Subgraph Centrality, Diversity Coefficient, Participation Coefficient, Within-Module Degree Z-Score

**Resilience** – network (in)vulnerability to insult. Properties include: Assortativity

# **Testing Relationships Between Graph Properties & Pathology**

Tools are available to calculate graph theoretic properties (e.g., Brain Connectivity Toolbox, Rubinov & Sporns, 2010). However, there is a dearth of tools available to test hypotheses about the relationship between graph properties and pathology in a (A) flexible and (B) valid manner. By flexible, we mean the ability to use of continuous and/or categorical predictors and between- and/or within-participant measures. By valid, we mean using non-parametric methods to calculate significance. Therefore, we developed the Graph Theoretic General Linear Model (GTG) MATLAB toolbox. The toolbox is freely available on NITRC: www.nitrc.org/projects/metalab\_gtg

#### References

Rubinov, M., Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. NeuroImage, 52, 1059-1069.

Freedman, D., Lane, D. (1983). A nonstochastic interpretation of reported significance levels. *Journal of Business* & Economic Statistics, 1, 292-298.

Power, J.D., Mitra, A., ... Petersen, S.E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. NeuroImage, 84, 320-341. Patel, A.X., Kundu, P., ... Bullmore, E.T. (2014). A wavelet method for modeling and despiking motion artifacts

from resting-state fMRI time series. *NeuroImage*, 95, 287-304. Jo, H.J., Gotts, S.J., ... Saad, Z.S. (2013). Effective preprocessing procedures virtually eliminate distance-

dependent motion artifacts in resting state FMRI. Journal of Applied Mathematics, 3, 339-352 Muschelli, J., Nebel, ... Mostofsky, S.H. (2014). Reduction of motion-related artifacts in resting state fMRI using aCompCor. NeuroImage, 96, 22-35.

Chen, G., Chen, G., ... Li, S. (2012). A method to determine the necessity for global signal regression in restingstate fMRI studies. *Magnetic Resonance in Medicine*, 68, 1828-1835. Saad, Z.S., Reynold, R.C., ... Cox, R.W. (2013). Correcting brain-wide correlation differences in resting-state fMRI. Brain Connectivity, 3, 339-352.

### **Stage 1 – Preprocessing**

This stage accepts raw fMRI 4D timeseries data, performs preprocessing, and extracts a processed timeseries for each input ROI. Processing includes standard options: slice-timing correction, motion correction, polynomial detrending. To remove motion variance, Power et al. (2014)'s motion scrubbing and Patel et al. (2014)'s wavelet despiking procedures are available. Several options are available for partialing nuisance signals, including mean global, white matter, and ventricular signals, and their (2<sup>nd</sup> order) 1<sup>st</sup> derivative. Partialing of motion correction parameters is available, along with the t-1 parameters, squared parameters, and the (2<sup>nd</sup> order) 1<sup>st</sup> derivatives. It is also possible to partial signal from only local white matter (within a 45mm sphere, J ventricular signal (5 components each) can be partialled instead of the Index (GNI) can be calculated and (automatically) used determine (2013)'s measure of average brain-wide correlations (GCOR) can be ca

## Stage 2 – Calculate Connectivity

This stage creates connectivity matrices (one per participant/repeated For block-design task fMRI, the toolbox will compute connectivity specified condition after dividing up the timeseries by condition. I HDR-related delay, timeseries are first deconvolved (using SPM's met) at actual onset/offset times. Detrending within each block is availab Partial correlation, Mutual information, & Robust (bendcorr) Correlat

#### **Stage 3 – Calculate Graph Properties**

This stage calculates graph theoretic properties for each participant repeated level using the Brain Connectivity Toolbox. The measure connectivity does not matter (e.g., correlation, diffusion tract strength)

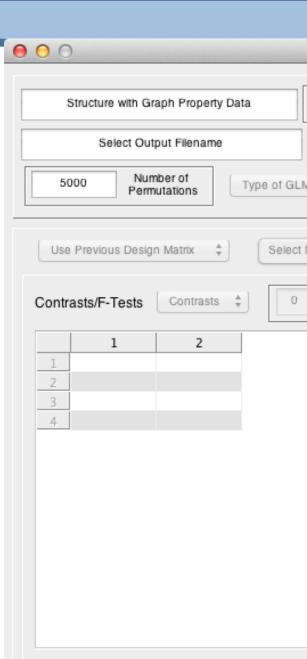
Properties are computed for both positive and negative weights from fully-connected and/or thresholded networks. For thresholde networks, properties are computed across a set of density threshold The minimum density is chosen such that the presence of disconnected

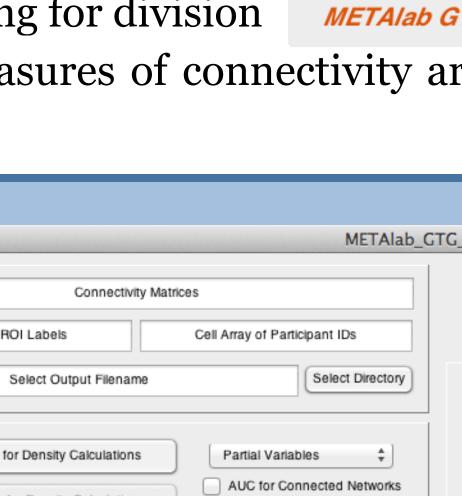
networks is not highly correlated with variables of interest. Specifical stratifying these variables. Mean networks are created for each (stratic remains connected is identified. This is done for each group (across network, and the maximum of these minima is chosen. An area under

This stage calculates GLMs on graph properties. Continuous & cat participant predictors and a categorical within-participant predictor a using a within-participant predictor, it is possible to test both main e X within interactions. It is possible to test the significance of indi contrasts between predictors, or F-tests across a set of predictors. Along with Ordinary Least Squares, it is possible to use Robust

Squares GLMs. Significance is determined via non-parametric permutation tests usi Freedman & Lane (1983) to deal with covariates.

#### Stage 4 – Run GLM





\varTheta 🔿 🔿 METAlab_GTG_prej	
	Processing Options
Cell Array of Participant IDs Cell Array of ROI Labels	Slice Timing Correction Slice Collection Order
Select Output Filename Select Directory	Detrending     Polynomial Order     +     Wavelet Despike
120     Enter # of Timepoints     2     Enter TR     Save Intermediate Files?	Bandpass Filter Type
Select Files	0.01     Highpass Cutoff     0.10     Lowpass Cutoff     0.3     FD Cutoff
Functional File Select	Butterworth Order
Select the 4D Functional File for the 1st Participant	Partialing Nuisance Signals
Select the 3D Brain Mask File for the 1st Participant	Motion Correction Parameters Global Signal
ROI Mask Select	t - 1 Parameters     1st Derivative of Signal       Squared Parameters     Test Necessity of Partialing
Select the 3D ROI Mask for the 1st Participant Image Space	1st Derivative of Parameters     Calculate GCOR
White Matter Mask     Select       Select the 3D White Matter Mask for the 1st Participant     Image Space	White Matter Signal     Use WM & Ventricular Principal Components       Scope of Mask
	1st Derivative of Signal     1st Derivative of Signal
Ventricle Mask     Select       Select the 3D Ventricle Mask for the 1st Participant     Image Space	
$\mathbf{T}_{\mathbf{r}} = \mathbf{r}_{\mathbf{r}} + $	METAIab GTG: Preprocessing Start
Jo et al., 2013). The first 5 principal components of white matter and ne mean (Muschelli et al., 2014). Chen et al. (2012)'s Global Negative whether it is necessary to partial global signal. Finally, Saad et al. alculated (useful as a GLM covariate).	
	METAlah CTC createconmate CIII
V Matrices	METAlab_GTG_createconmats_GUI
	Series Data Set Output Filename Select Directory
condition).	al Detrending Order
y matrices for each user-	g Block Timing Info OR Input Block Timing Info Manually
n order to compensate for	
hod), allowing for division <b>METAIAD G</b>	TG: Create Connectivity Matrices Start
le. Four measures of connectivity ar	e available: Pearson correlation,
tion.	
⊖ O METAlab_GTG_	_calcproperties_GUI
O METAlab_GTG_ Connectivity Matrices	10000 Number of Runs for Calculate Max Club Size
Connectivity Matrices Cell Array of ROI Labels Cell Array of Participant IDs	Number of Runs for Calculate Max Club Size
Connectivity Matrices Cell Array of ROI Labels Cell Array of Participant IDs Select Output Filename Select Directory	10000       Number of Runs for Modularity       Calculate Max Club Size         30       Max Club Size         Use Absolute Value of Weights       Max Club Size         Properties for Fully Connected Matrices       Properties for Thresholded Matrices         None       None
Connectivity Matrices Cell Array of ROI Labels Cell Array of Participant IDs Select Output Filename Select Directory	10000       Number of Runs for Modularity       Calculate Max Club Size         30       Max Club Size         30       Max Club Size         Very Properties for Fully Connected Matrices       Properties for Thresholded Matrices         None       Assortativity Characteristic Path Length Clustering Coefficient Diversity Coefficient
Connectivity Matrices Cell Array of ROI Labels Cell Array of Participant IDs Select Output Filename Select Directory Select Variables for Density Calculations Partial Variables \$ AUC for Connected Networks	10000       Number of Runs for Modularity       Calculate Max Club Size         30       Max Club Size         30       Max Club Size         Vise Absolute Value of Weights       Properties for Thresholded Matrices         Properties for Fully Connected Matrices       Properties for Thresholded Matrices         None       Assortativity         Characteristic Path Length       Clustering Coefficient         Diversity Coefficient       Degree         Edge Betweeness Centrality       Density         Eigenvector Centrality       Edge Betweeness Centrality         Edge Betweeness Centrality       Edge Netweeness Centrality
t/ of ). Select Variables for Density Calculations Select Covariates for Density Calculations Max Density for Max Density for Max Density for Max Density for Max Density for	1000       Number of Runs for Modularity       Calculate Max Club Size         30       Max Club Size         0       Properties for Thresholded Matrices         None       Assortativity         Characteristic Path Length       Clustering Coefficient         Diversity Coefficient       Degree         Edge Betweeness Centrality       Density         Global Efficiency       Elgenvector Centrality         Local Efficiency       Global Efficiency         Node Betweeness Centrality       Cocal Efficiency         Node Betweeness Centrality       Cocal Efficiency         Node Strength       Matching Index
t/ of ). Max Density for Max Density for Matrices	10000       Number of Runs for Modularity       Calculate Max Club Size         30       Max Club Size         Properties for Fully Connected Matrices       Properties for Thresholded Matrices         None       Assortativity Characteristic Path Length Clustering Coefficient Diversity Coefficient       Properties for Thresholded Matrices         None       Assortativity Characteristic Path Length Clustering Coefficient Degree       Degree         Local Efficiency       Degree       Density         Local Efficiency       Edge Betweeness Centrality Eigenvector Centrality Global Efficiency       Edge Betweeness Centrality Local Efficiency         Node Betweeness Centrality       K-Coreness Centrality Local Efficiency       Matching Index Node Betweeness Centrality         PageRank Centrality       PageRank Centrality       PageRank Centrality         Participation Coefficient       PageRank Centrality       PageRank Centrality         Participation Coefficient       Ethicipent       Participation Coefficient         Rich Club Networks       Transtivity       <
t/ of ). Max Density for Positive Weight Matrices Negative Weight Matrices	10000       Number of Runs for Modularity       Calculate Max Club Size         30       Max Club Size         Properties for Fully Connected Matrices       Properties for Thresholded Matrices         None       Assortativity         Characteristic Path Length       Clustering Coefficient         Diversity Coefficient       Degree         Eigenvector Centrality       Eigenvector Centrality         Global Efficiency       Eigenvector Centrality         Matching Index       Node Betweeness Centrality         Node Strength       Calculate Max Club Networks         Participation Coefficient       Node Betweeness Centrality         PageRank Centrality       PageRank Centrality         PageRank Centrality       PageRank Centrality         PageRank Centrality       Participation Coefficient         Node Betweeness Centrality       PageRank Centrality         PageRank Centrality       Participation Coefficient
t/ of ). Max Density for Natrices 0.01 Density Step for Positive Weight 0.01 Density Step for 0.01 Density Step for	10000       Number of Runs for Modularity       Calculate Max Club Size         30       Max Club Size         30       Max Club Size         Properties for Fully Connected Matrices       Properties for Thresholded Matrices         None       Assortativity         Characteristic Path Length Clustering Coefficient       Characteristic Path Length         Diversity Coefficient       Degree         Edge Betweeness Centrality       Elgenvector Centrality         Elgenvector Centrality       Elgenvector Centrality         Global Efficiency       Matching Index         Node Betweeness Centrality       Node Betweeness Centrality         PageRank Centrality       PageRank Centrality         Page
t/ of ). Max Density for N. Matrices d. METTALab GTG: Calculate Properties Cell Array of Participant IDs Cell Array of POI Labels Cell Array of Participant IDs Cell Array of Participant IDs Select Output Filename Select Directory Select Variables for Density Calculations Partial Variables Partial Variables Partia Par	10000       Number of Runs for Modularity         Image: Description of the state
t/ of ). Metables for Density Calculations ed s. d METAIAb GTG: Calculate Properties Variables of int	10000       Number of Runs for Modularity         0       Use Absolute Value of Weights         Properties for Fully Connected Matrices         Properties for Fully Connected Matrices         None         Assortativity Characteristic Path Length Clustering Coefficient Diversity Coefficient Rich Cub Networks Transitivity Within-Module Degree Z-Score       Properties for Thresholded Matrices         None       None       None         Participation Coefficient Transitivity Within-Module Degree Z-Score       Properties for Thresholded Matrices         None       None       None         Participation Coefficient Transitivity Within-Module Degree Z-Score       None         Subgraph Centrality Participation Coefficient Rich Cub Networks Subgraph Centrality Transitivity Within-Module Degree Z-Score         Subgraph Centrality Within-Module Degree Z-Score       Subgraph Centrality Transitivity Within-Module Degree Z-Score
t/ of ). med ed s. d METTAIab GTG: Calculate Properties 10, the user specifies variables of int atification) group, and the minimum	10000       Number of Runs for Modularity         Use Absolute Value of Weights         Properties for Fully Connected Matrices         None         Assontativity Characteristic Path Length Clustering Coefficient Edge Betweeness Contrality Elgenvector Centrality Global Efficiency Matching Index Node Betweeness Centrality Elgenvector Centrality RegeRank Centrality PageRank Centrality Participation Coefficient Rich Cub Networks Transitivity Within-Module Degree Z-Score         Terrest, and groups are created by n density at which that network
t/ of ). m ed s. d METAIab GTG: Calculate Properties METAIable, across all selected variable, across all selected variables of interval action	1000       Number of Runs for Modularity         Use Absolute Value of Weights         Properties for Fully Connected Matrices         Assortativity Characteristic Path Length Clustering Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Participation Coefficient         Participation Coefficient         Participation Coefficient         Rich Cub Networks         Transitivity         Within-Module Degree Z-Score         Within-Module Degree Z-Score         Within-Module Degree Z-Score
t/ of ). m ed s. d METAIab GTG: Calculate Properties METAIable, across all selected variable, across all selected variables of interval action	1000       Number of Runs for Modularity         Use Absolute Value of Weights         Properties for Fully Connected Matrices         Assortativity Characteristic Path Length Clustering Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Participation Coefficient         Participation Coefficient         Participation Coefficient         Rich Cub Networks         Transitivity         Within-Module Degree Z-Score         Within-Module Degree Z-Score         Within-Module Degree Z-Score
t/ of ). m ed s. d METAIab GTG: Calculate Properties METAIable, across all selected variable, across all selected variables of interval action	1000       Number of Runs for Modularity         Use Absolute Value of Weights         Properties for Fully Connected Matrices         Assortativity Characteristic Path Length Clustering Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Diversity Coefficient         Participation Coefficient         Participation Coefficient         Participation Coefficient         Rich Cub Networks         Transitivity         Within-Module Degree Z-Score         Within-Module Degree Z-Score         Within-Module Degree Z-Score
t/ of ). med ds. d 1. ty, the user specifies variables of int atification) group, and the minimum each variable, across all selected var the curve (AUC) across density levels	1000       Number of Huns for Modularity         Use Absolute Value of Weights         Properties for Fully Connected Matrices         None         Mascriativity         Characteristic Path Length         Custering Coefficient         Properties for Fully Connected Matrices         None         Mascriativity         Characteristic Path Length         Custering Coefficient         Properties for Thresholded Matrices         None         Vite State Contrainty         Within-Module Degree Z-Score
t/ of ). m ed s. d	10000       Number of Runs for Modulany         0       Use Absolute Value of Weights         Properties for Fully Connected Matrices         Number of Runs for Modulany         Operties for Fully Connected Matrices         Number of Runs for Modulany         Properties for Fully Connected Matrices         Number of Runs for Modulany         Properties for Fully Connected Matrices         Number of Runs for Modulany         Properties for Fully Connected Matrices         Number of Runs for Modulany         Properties for Fully Connected Matrices         Number of Runs for Modulany         Properties for Fully Connected Matrices         Number of Runs for Modulany         Properties for Thresholded Matrices         Number of Runs for Modulany         Properties for Thresholded Matrices         Number of Runs for Matching Index         Num
t/ of ). Select Variables for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Binarce Thresolded Matrices Binarce Thresolded Matrices Matrices Start Start Start Start Start Start Start Start Start Start Start Select Covariable, accross all selected var the curve (AUCC) across density levels Select Coupt Flename Select Coupt Flename	10000       Number of Runs for Modularity <ul> <li>Calculate Max Club Size</li> <li>Max Club Size</li> <li>Max Club Size</li> </ul> Use Absolute Value of Weights               Properties for Thresholded Matrices                 Properties for Fully Connected Matrices               Properties for Thresholded Matrices                 Properties for Fully Connected Matrices               Properties for Thresholded Matrices                 Prost Max               Prost                 Prosteries               Centrainy                 Properties                Prostel
t/ of ). med s. d the d s. d the d s. d the transies tor Density Calculations Select Variables for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Net AuC for Connected Networks Binarize Thresholded Matrices Binarize Thresholded Matrices Matrices Metrafiab GTG: Calculate Properties Start Uly, the user specifies variables of intra atification) group, and the minimum each variable, across all selected variations the curve (AUCC) across density levels the curve (AUCC) across density levels select Output Flename correct Curve Flename correct Flename correct Curve Flename correct Curve Flename correct Curve Flename correct Flename correct Curve Flename correct Curve Flename correct Flename correc	10000       Number of Fluns for Modularity       30       Max Club Size         30       Max Club Size       30       Max Club Size         Properties for Fully Connected Matrices         Properties for Fully Connected Matrices       Properties for Thresholded Matrices         Participation Contrainty       Clustering Coefficient         Coal Efficienty       Clustering Coefficient         Rich Club Networks       Strainty         Participation Coefficient       Rich Club Networks         Small Within-Module Degree Z-Score       Small Word Bestweeness Centrainty         Participation Coefficient       Rich Club Networks         Small Within-Module Degree Z-Score       Small Word Networks         Small Word Networks       Small Word Networks         Small Word Networks       Small Mord Networks         Score
t/ of ). m ed select Variables tor Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Binarize Thresholded Matrices Used Select Covariates for Density Calculations Select Covariates for Density Calculations O.6 Positive O.6 Negative Matrices Matrices Matrices Metralab GTG: Calculate Properties Start Illy, the user specifies variables of intra atification) group, and the minimum each variable, across all selected var the curve (AUC) across density levels the curve (AUC) across density levels select Curve for Property Data is select Curve for Property Data Select Curve for Property Data (Select Curve for Property Data) Select Curve for Property Data (Select Curve for Property Data) Select Curve for Property Data (Select Curve for Property Data) (Select Curve for Property Data)	10000       Number of Huns for Modularity         10000       Number of Huns for Modularity         10000       Number of Huns for Max Club Size         10000       Number of Huns for Hundes       Number of Huns for Max Club Size         10000       Number of Huns for Hundes       Number of Huns for Hundes         10000       Number of Huns for Hundes       Number of Huns for Hundes         10000       Number of Huns for Hundes       Number of Huns for Hundes         10000       Number of Hundes       Number of Hundes         10000       Number of Hundes       Numer of Hundes <td< th=""></td<>
t/ of ). m ed select Variables for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Binarize Thresholded Matrices Binarize Thresholded Matrices Matrices Matrices Matrices Matrices Matrices Matrices Matrices Start Illy, the user specifies variables of intra tatification) group, and the minimum each variable, across all selected var the curve (AUC) across density levels select Output Fienare tegorical between- the accepted. When iffects and between	10000       Number of Huns for Modularity         1       Use Absolute Value of Weights         Properties for Fully Connected Matrices         Associativity Associativity Contacterestic Path Length Contacterestic Path Length Conterestic Path Length Contacterestic Path Le
t/ of ). m ed select Variables for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Select Covariates for Density Calculations Binarce Thresholded Matrices Binarce Thresholded Matrices Start Use Percent Beage Matrices Binarce Thresholded Matrices B	10000       Number of Huns for Modularity         10000       Number of Huns for Modularity         10000       Number of Huns for Max Club Size         10000       Number of Huns for Hundes         10000       Number of Hundes         10000       Number of Hundes         10000       Number of Hundes <td< th=""></td<>
t/ of	10000       Number of Huns for Modularity         10000       Number of Huns for Modularity         10000       Number of Huns for Max Club Size         10000       Number of Huns for Hundes         10000       Number of Hundes         10000       Number of Hundes         10000       Number of Hundes <td< th=""></td<>
t/ of	10000       Number of Huns for Modularity         10000       Number of Huns for Modularity         10000       Number of Huns for Max Club Size         10000       Number of Huns for Hundes         10000       Number of Hundes         10000       Number of Hundes         10000       Number of Hundes <td< th=""></td<>
t/ of select Output Flename selected watcos select Dectory select Output Flename select Dectory select Output Flename selected watcos select Dectory select Covariates for Density Calculations bracks thresholded Matrices to Density Calculations bracks thresholded Matrices and the user specifies variables of intra attification) group, and the minimum each variable, across all selected var the curve (AUC) across density levels the curve (AUC) across density levels for levels and between fifects and between ividual predictors, or Trimmed Least	10000       Number of Huns for Modularity         10000       Number of Huns for Modularity         10000       Number of Huns for Max Club Size         10000       Number of Huns for Hundes         10000       Number of Hundes         10000       Number of Hundes         10000       Number of Hundes <td< th=""></td<>
t/ of	10000       Number of Huns for Modularity         10000       Number of Huns for Modularity         10000       Number of Huns for Max Club Size         10000       Number of Huns for Hundes         10000       Number of Hundes         10000       Number of Hundes         10000       Number of Hundes <td< th=""></td<>



