**First Initial Data mining Work**

**“*how abundant are events with time ?*”**

**Goal:**

**1.** Ability to “count events and zone/pattern boundaries” in selected or merged datapacks within sliding-window “bins” (1/2 myr bins or other subdivision). We want to find evolution-facies changes in the past through this crude reflection.

**2.** Ability to show the frequency of events in our TSCreator software

**Method:** Load ALL the databases (ca. 500,000 entries at this point), then exclude the “curve” type. *How many events are occurring in each 1-million-year window for the past 600 million years?*

1. Convert column data to MySQL table

Tscreator java software uses Hashmap for saving the datapacks. We have a converter sub-program which will be used to create MySQL tables to represent columns.

b) Analyze/ Data Mine using Python scripts

Having a MySQL table for all the columns will be easier to parse using Python scripts (and SQL queries) to slide over million year windows to find the frequency of events. All other future data mining task will also be done simple. For research publication, generating plots will be easier.

Clarification Needed:

1. Which events to extract?
2. How to merge multiple datapacks?
3. Data format in the SQL table? etc..

c) Interface to main TSCreator program to show the data mining task results

Using python scripts or MySQL table dumping feature, again it’s possible to create datapack columns and integrate with the existing datapacks in the preferable visualization format to finally visualize the answers to the data mining questions we will be exploring

**Main Questions to be explored:**

(a) What intervals are relative upheavals in Earth’s history and which intervals are anomalously “quiet”?

(b) Do these oscillations support the “Nemesis” or other periodic catastrophe models (ca. 20 to 60 million-year periods if one does an evolutive spectra analysis?),

(c) Are there suggestions of the postulated long-term (ca. 2 to 5 myr) astronomical cycles, and (d) which peaks in events correlate with peaks in volcanic activity, continental collisions and/or major asteroid impacts?

(d) What happens if we run 2 or 5 million year sliding windows instead?

(e) What would be the projected main changes for the next million years (our own future) based on any periodicities and trends?

**Papers to read:**

1. [Nemesis reconsidered](https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/mnrasl/407/1/10.1111/j.1745-3933.2010.00913.x/3/407-1-L99.pdf?Expires=1500425887&Signature=eUrVytQOnAs8MfmAQFCqx8hfi0xK1TwFZNlG1BwRJt6DxyyiTzba7d0CpX7mgFuNaGqOm3lmQWZYb9QQK3DRcrlOMIRD8qV--MENqUNBh9BWKPtcojpm8k-V~v70O46C7EGMK9-YTaFvgm63x0T7axAxiKCSQvogSfmcvuNi9GHb-nf8QC4wepdQSX9Eam~a5TEnUShja-8G5qLdV48YCRnOoh88P1aPV8lBcsRcf9FC8viNmQjXq787Oc2Z-mW58WtJ~Cb66Q2L5c4iwrVESrds9jYeoKsIzmrC7IXpNLe0EZBbW05tYvja2LN2hLJ6fu4PYnUtK19Byt8MUSOZ-A__&Key-Pair-Id=APKAIUCZBIA4LVPAVW3Q)
2. [Periodicity in extinction and the problem of catastrophism in the history of life](http://jgs.geoscienceworld.org.ezproxy.lib.purdue.edu/content/jgs/146/1/7.full.pdf)
3. [Large period astronomical cycles from the Triassic to Jurassic bedded chert sequence (Inuyama, Japan); Geologic evidences for the chaotic behavior of solar planets.](https://link.springer.com/content/pdf/10.5047%2Feps.2012.09.004.pdf)
4. [Comparing the geological and fossil records: implications for biodiversity studies.](https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/mnrasl/407/1/10.1111/j.1745-3933.2010.00913.x/3/407-1-L99.pdf?Expires=1500425887&Signature=eUrVytQOnAs8MfmAQFCqx8hfi0xK1TwFZNlG1BwRJt6DxyyiTzba7d0CpX7mgFuNaGqOm3lmQWZYb9QQK3DRcrlOMIRD8qV--MENqUNBh9BWKPtcojpm8k-V~v70O46C7EGMK9-YTaFvgm63x0T7axAxiKCSQvogSfmcvuNi9GHb-nf8QC4wepdQSX9Eam~a5TEnUShja-8G5qLdV48YCRnOoh88P1aPV8lBcsRcf9FC8viNmQjXq787Oc2Z-mW58WtJ~Cb66Q2L5c4iwrVESrds9jYeoKsIzmrC7IXpNLe0EZBbW05tYvja2LN2hLJ6fu4PYnUtK19Byt8MUSOZ-A__&Key-Pair-Id=APKAIUCZBIA4LVPAVW3Q)

Questions to be explored - Effect of Permo-Triassic extinction

1) Why the end of the Permian provided opportunities for the mollusks?

2) Fish evolution in the upper Permian, it seems to be the least well-known interval for fish interval.

3) Do some investigation on the deep ocean rocks (deep sea cherts) in native Japan.

4) What’s the cause of severe radiolarian extinction at the end of the Permian? Radiolarian chert.. Why did the radiolarians survived the Perms-Triassic crisis in the Arrow rocks in Newzealand ?

5) Verify the shift from Eukaryotic green algae(Organic matter) to cyanobacteria (primitive-type of photosynthesizing bacterial group).

6) Why was there green sulfur bacteria at that time? Correlation with anoxia, prevalence of hydrogen sulfide.

7) Check the life on land - Extermination of dinocephalians, evolution of dicynodonts evolving into plant-eating creature, gorgonopisans as large-predator, pareiasurds as big, bulky reptiles with lumps and horns, amphibians(temnospnondils) to the evolution of crocodiles.

8) Check whether we have data for Karoo region of South Africa and the foothills of the Urals of Russia,- Verify the extinction of terrestrial animals

9) Is it true that Greenland area had healthy, thriving forest ? Can we incorporate the data on spore and pollens which provide the information about flora and fauna during perms-triassic from Cindy Looy’s research.

10) Check the coal formation which was the result of Permo-Triassic extinction. Coals are the product of lush forests growing in swaps. Check Antarctica and Australian (high southern latitudes) data. Why did coal formation abruptly stop at the end of Permian.

11) Do we have data for Reduviasporonites (spore) in equatorial Perms-Triassic boundary rocks. Is pond-scum hypothesis (freshwater algae which create pond-scum was proliferated after the mass extinction and the spores were from these algae ) wrong ?

12) The effect on sedimentation due to the death of plants, trees? The change of river style from meandering to braided, causality or correlation with the extinction.

**Goal**: We can search over our dataset to find potential oil/gas reservoir which are not already detected. Initially we can look at the existing oil/gas reservoir and find the patterns, activities, anomalies in other columns.