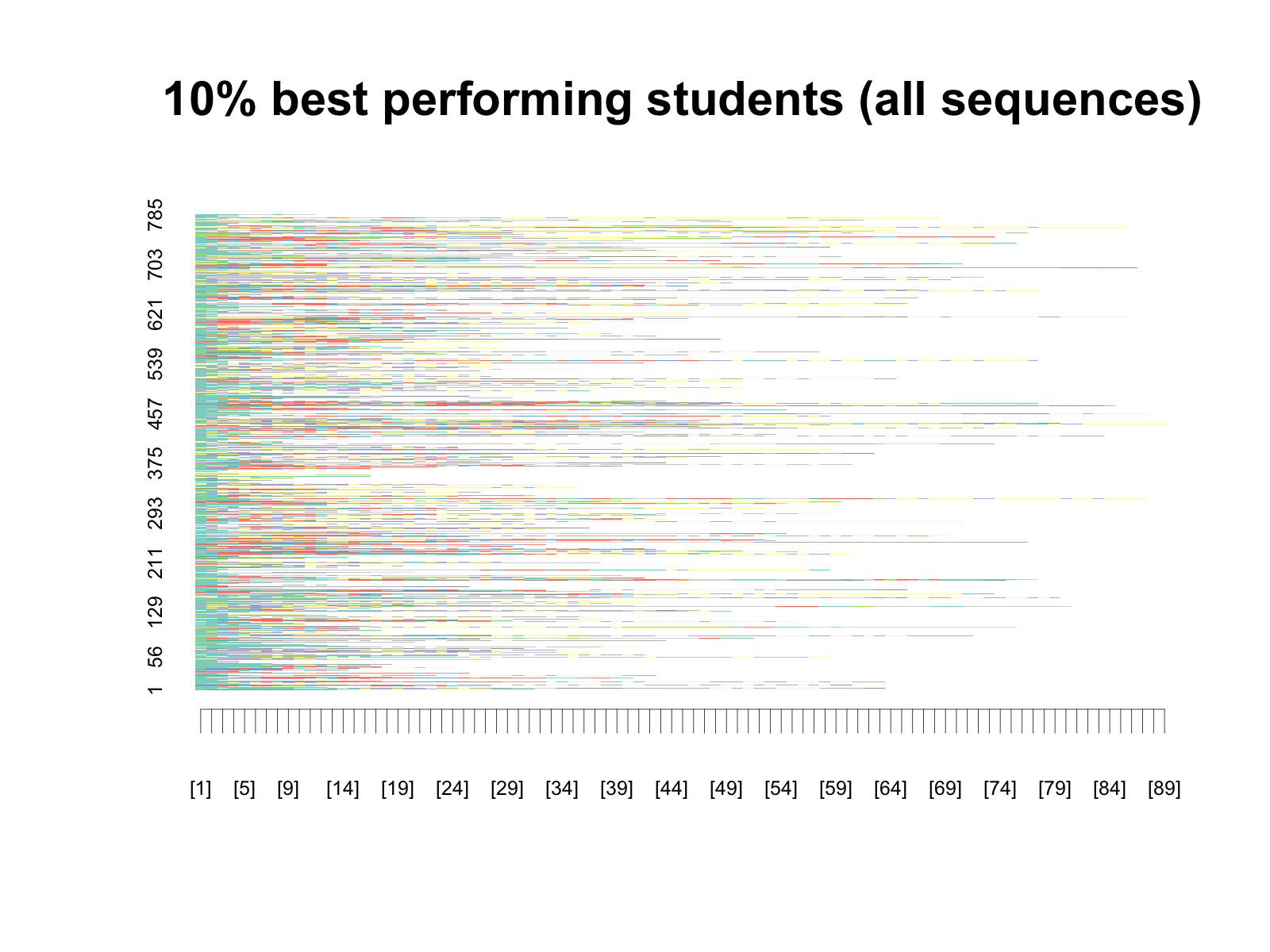
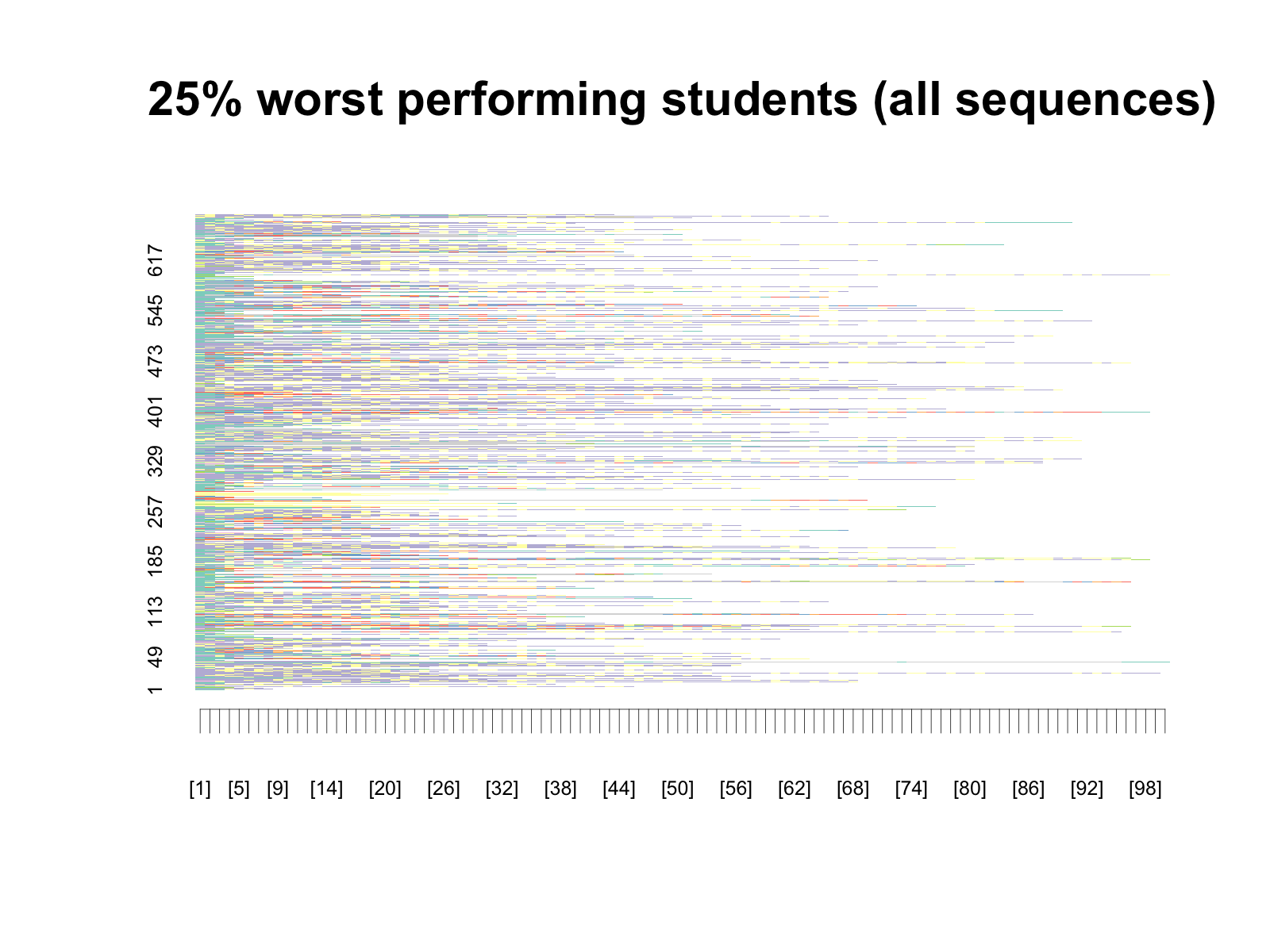
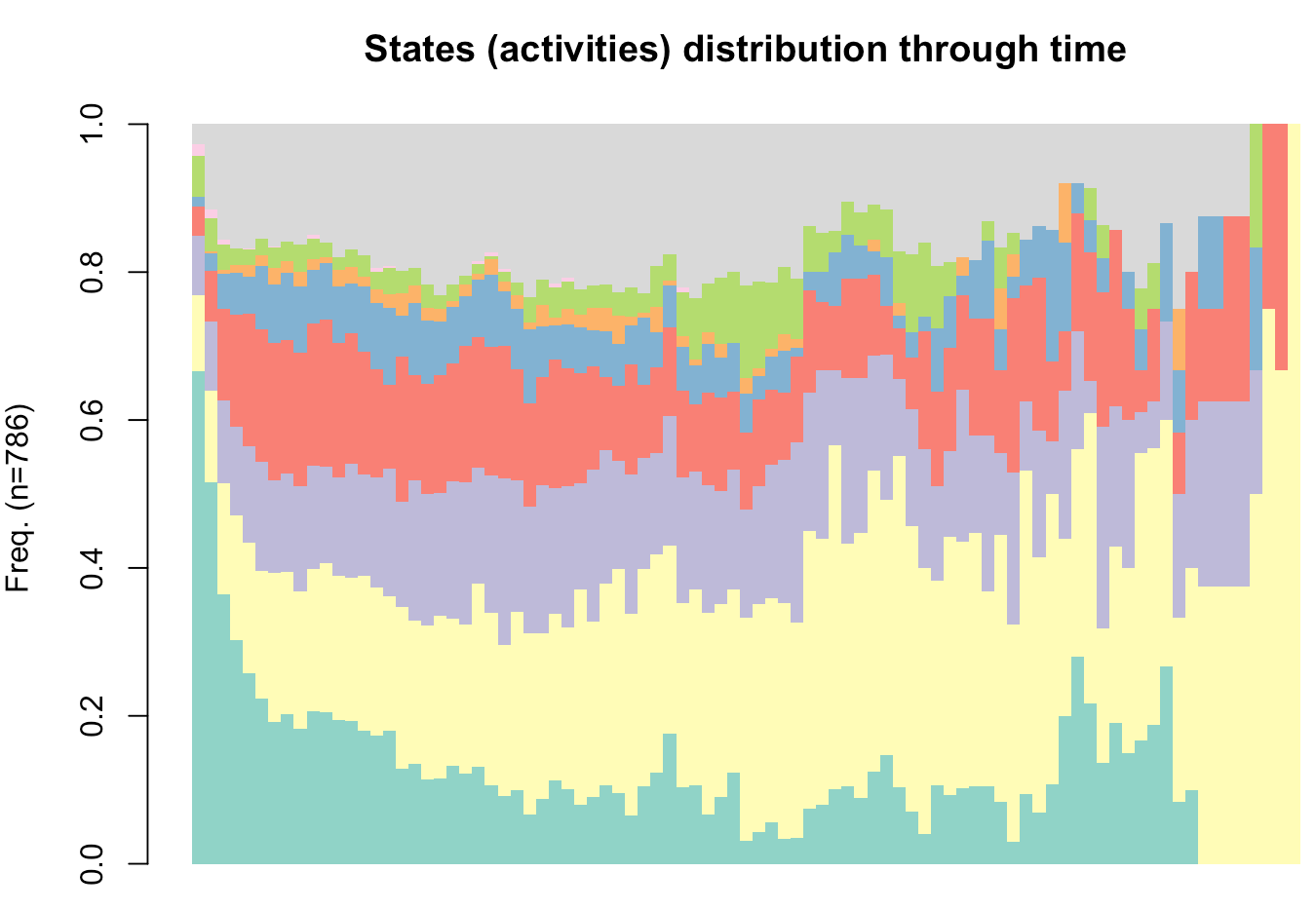
ALL SEQUENCES OF THE TOP 10% PERFORMING STUDENTS



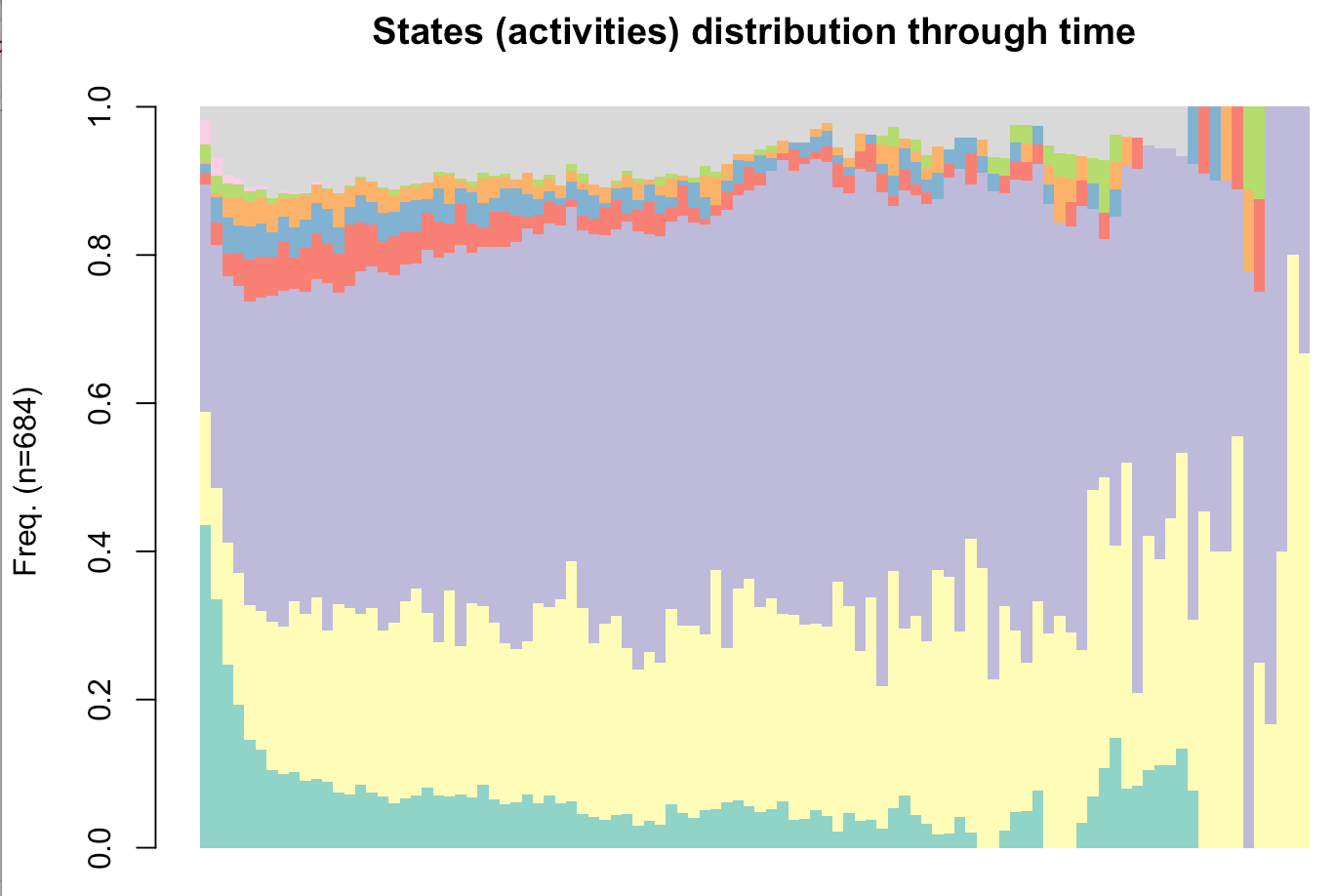
ALL SEQUENCES OF THE 25% WORST PERFORMING STUDENTS



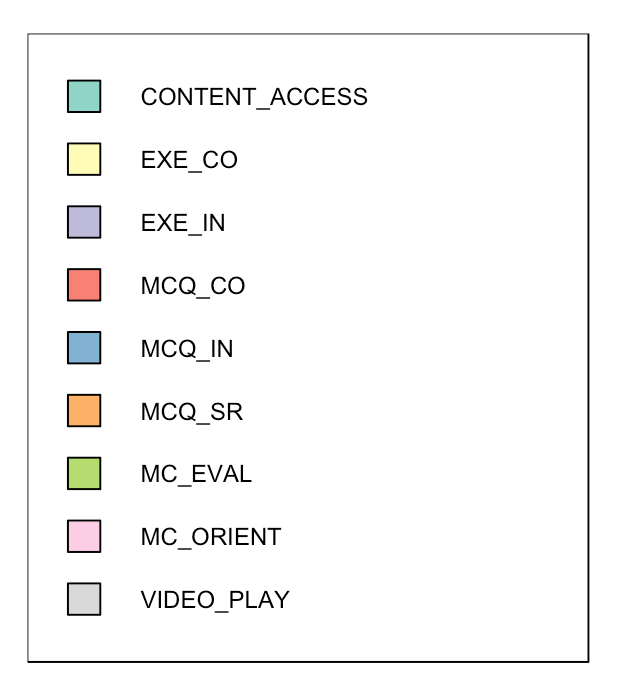
STATE (ACTIVITY) DISTRIBUTION ALONG SEQUENCES’ TIME POINTS FOR THE 10% BEST PERFORMING STUDENTS



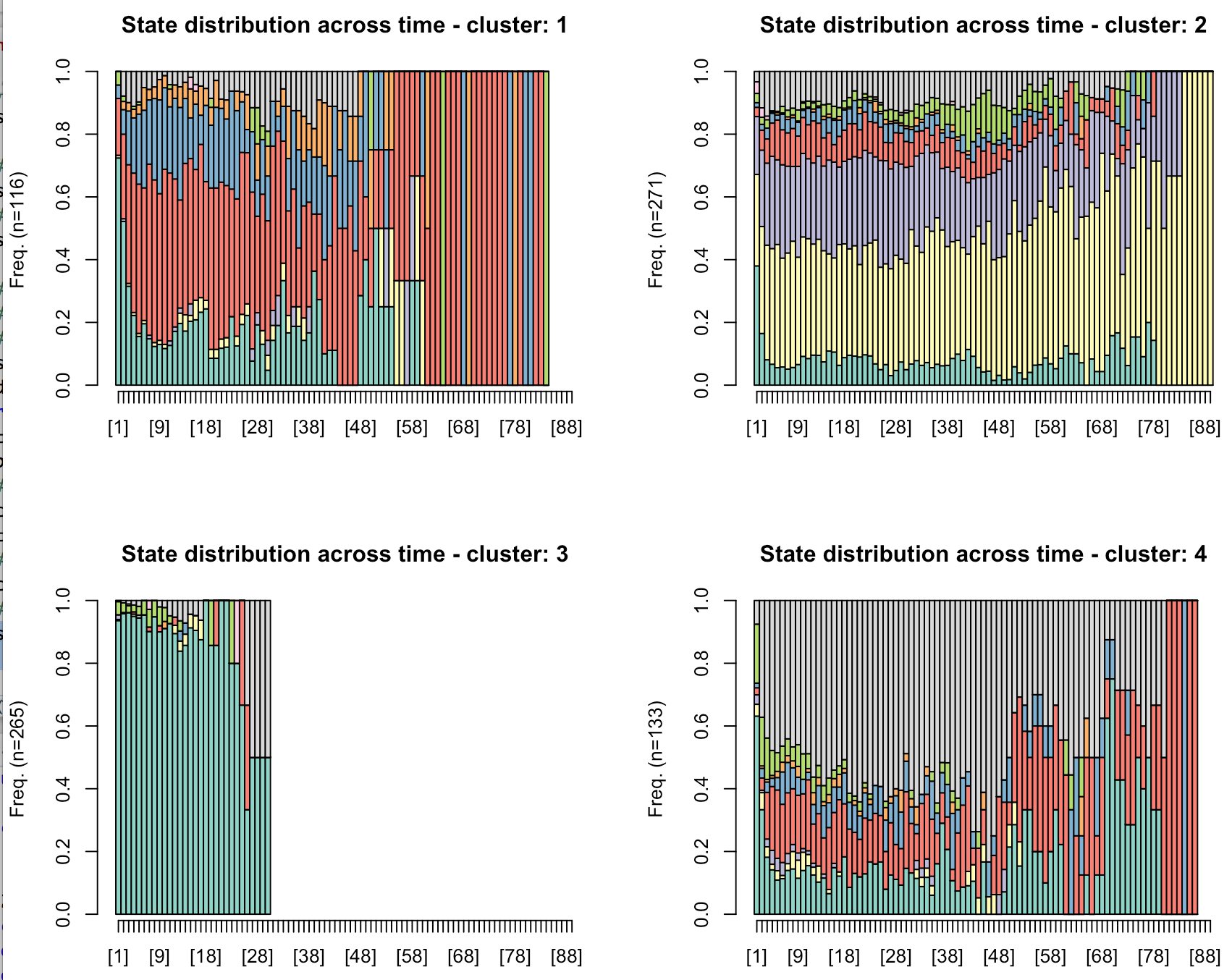
THE SAME DIAGRAM FOR 25% WORST PERFORMING STUDENTS



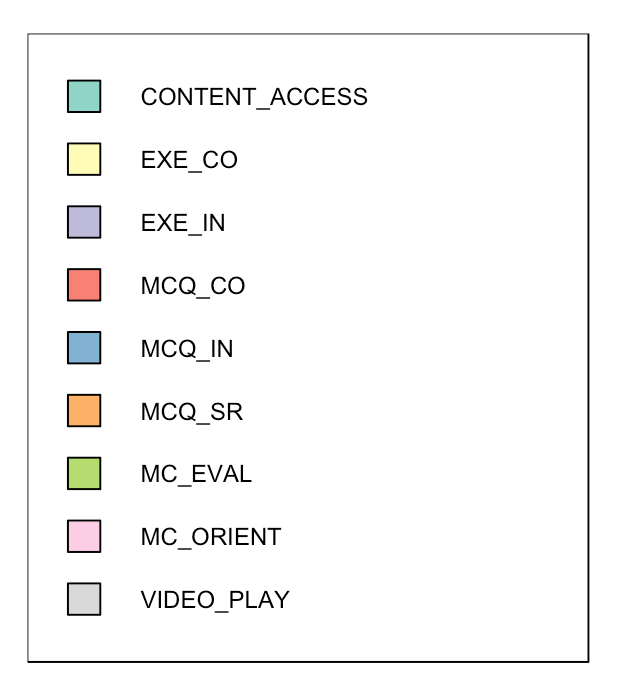
Legend:



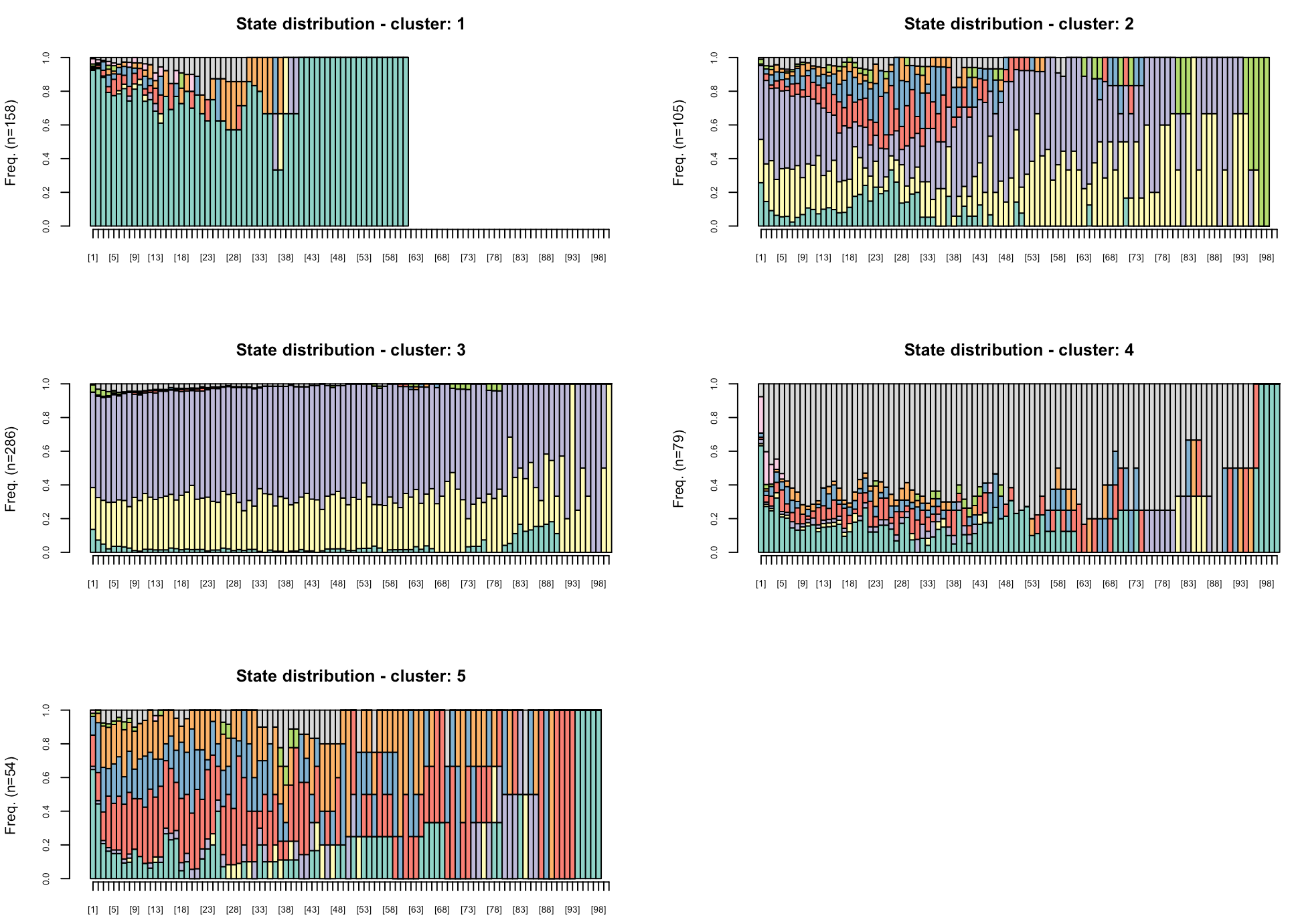
SEQUENCE CLUSTERS OF THE TOP 10% PERFORMING STUDENTS



Legend:



SEQUENCE CLUSTERS OF THE 25% WORST PERFORMING STUDENTS



SEQUENCES OF STUDENT GROUPS OBTAINED THROUGH LCA CLUSTERING

The following figures show distribution of activities (states) along sequences’ time points for the 5 student groups/classes obtained by applying the LCA method on weekly clusters.

For each LCA group, I’ve extracted learning sessions performed by the students assigned to that group, and then created sequences out of the session data. The figures suggest some differences in the groups’ learning sessions; though, I still haven’t found a way to objectively compare the groups’ sequences. I’ve tried with some sequence-related statistics (given in the table below), but that proved non-informative.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

Some sequence-related statistics, computed for LCA classes

======= ========= ======= ================ ========== ==============

\ num.users num.seq avg.seq.per.user seq.length seq.entropy

======= ========= ======= ================ ========== ==============

class\_1 92 2953 32.10 16 (5, 34) 0.28 (0, 0.36)

class\_2 59 2573 43.61 12 (4, 28) 0.29 (0, 0.48)

class\_3 59 1353 22.93 10 (4, 24) 0.27 (0, 0.41)

class\_4 44 2353 53.48 11 (4, 29) 0.30 (0, 0.51)

class\_5 36 2124 59.00 11 (4, 31) 0.31 (0, 0.56)

======= ========= ======= ================ ========== ==============

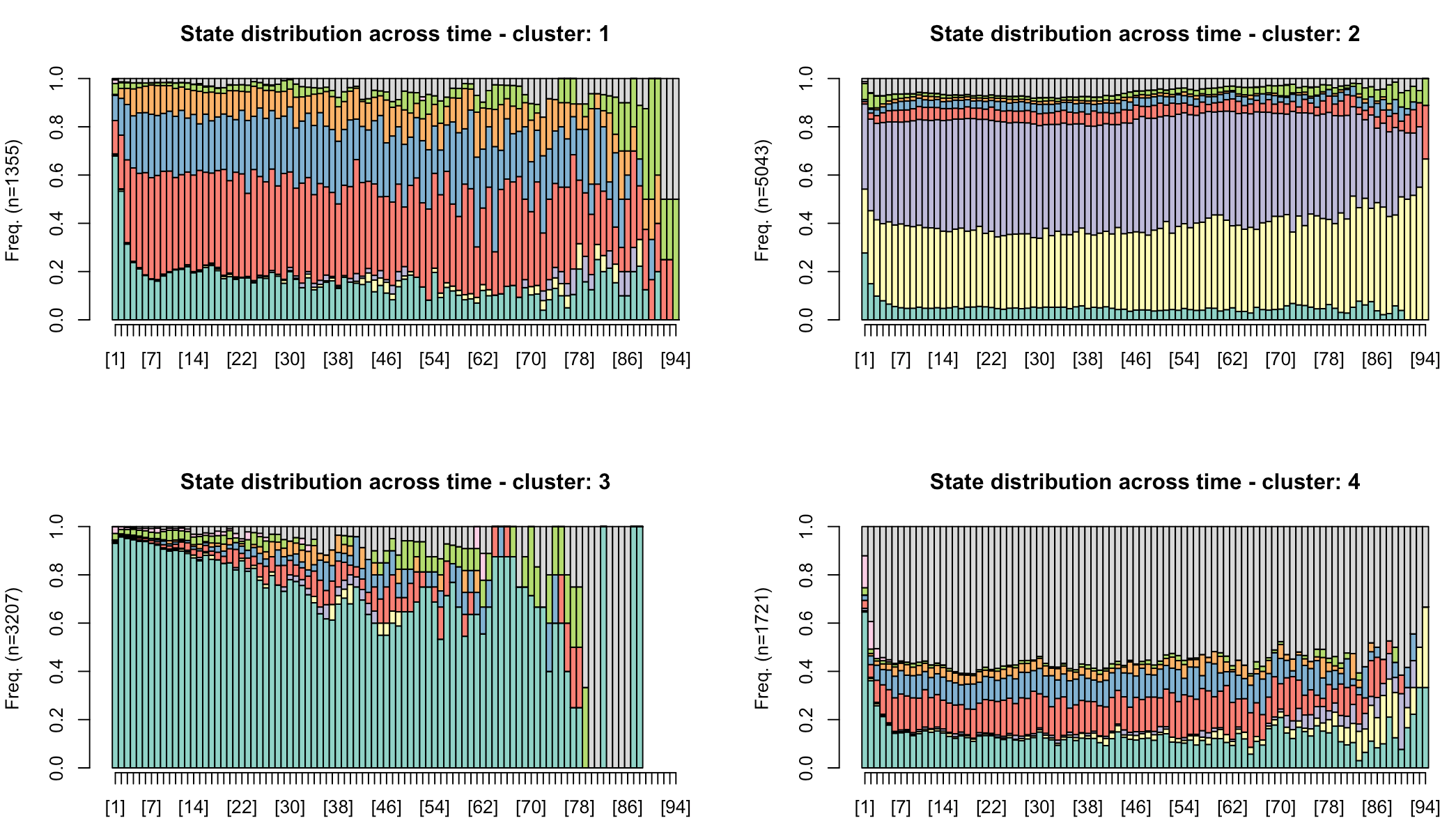
SEQUENCE ANALYSIS OF LEARNING TRACES OF ALL THE STUDENTS

Quick summary of the analysis that led to the results shown below:

Learning sessions, for weeks 2-13, of all the students (n=290) were used as the input for the analysis. First, learning sessions data were transformed into sequences (as required for the analysis with TraMineR). Then, the sequences (n=11,326) were clustered, using hierarchical clustering method (Ward). The solution with 4 clusters was chosen as the best one (based on the dendogram).

The next step was to cluster students based on the number of sequences (learning sessions) per sequence clusters. In particular, the clustering of sequences produced for each student 4 variables, each one having as its value the number of the student’s sequences in one of the 4 identified clusters. These 4 variables were then used to cluster the students (again using hierarchical clustering).

**Step 1: Clusters of sequences (learning sessions)**



Legend:

|  |  |
| --- | --- |
|  |  |

**Step 2: Clustering of students based on the sequence (learning session) clusters**

Solution with 4 clusters was chosen as the best

Comparison of clusters with respect to the features used for clustering: sequence clusters, i.e., clusters of learning sessions

========== ================ ================ ================ ================

attributes Median; (Q1, Q3) Median; (Q1, Q3) Median; (Q1, Q3) Median; (Q1, Q3)

========== ================ ================ ================ ================

cluster 1 2 3 4

freq 22 70 133 65

cl1 12.5; (11, 15) 6.5; (4, 10) 3; (1, 5) 1; (0, 3)

cl2 26; (23.25, 29) 21; (18, 23) 18; (16, 21) 10; (7, 13)

cl3 29.5; (24, 33.5) 18; (15, 21) 8; (5, 11) 3; (2, 5)

cl4 12; (9, 13.75) 9; (7, 12) 4; (2, 7) 1; (0, 2)

========== ================ ================ ================ ================

Basic statistics for *midterm exam score*

===== === ====== === ===

class N median Q1 Q3

===== === ====== === ===

1 22 15 13 17

2 70 16 13 17

3 133 13 11 16

4 65 11 8 15

===== === ====== === ===

Comparison of clusters based on the midterm exam score, with FDR correction applied

=== === === ======= ======== ====== ===========

\ c1 c2 Z p r significant

=== === === ======= ======== ====== ===========

5 2 4 5.2789 0.000000 0.4543 YES

3 1 4 3.8533 0.000068 0.4131 YES

4 2 3 3.4314 0.000546 0.2408 YES

6 3 4 3.3489 0.000741 0.2380 YES

2 1 3 2.3378 0.018746 0.1878 YES

1 1 2 -0.0322 0.976299 0.0034 NO

=== === === ======= ======== ====== ===========

Basic statistics for *final exam score*

===== === ====== ===== =====

class N median Q1 Q3

===== === ====== ===== =====

1 22 22.5 17.25 29.75

2 70 21.0 15.25 30.75

3 133 16.0 13.00 24.00

4 65 14.0 10.00 18.00

===== === ====== ===== =====

Comparison of clusters based on the final exam score, with FDR correction applied

=== === === ====== ======== ====== ===========

\ c1 c2 Z p r significant

=== === === ====== ======== ====== ===========

5 2 4 5.2587 0.000000 0.4526 YES

3 1 4 4.0821 0.000022 0.4376 YES

4 2 3 3.3078 0.000868 0.2322 YES

6 3 4 3.2844 0.000942 0.2334 YES

2 1 3 2.4951 0.011951 0.2004 YES

1 1 2 0.1741 0.864580 0.0182 NO

=== === === ====== ======== ====== ===========

In summary, all the cluster pairs, except for the cluster 1 - 2 pair, are significantly different (even after applying the FDR correction for multiple testing) in terms of both midterm exam score and final exam score.

**Summary of the findings**

If we refer to the 4 sequence clusters (i.e., clusters of learning sessions) identified in Step 1 as learning patterns or learning strategies, those patterns/strategies might be described as:

* Pattern/strategy 1: focus on formative assessment, followed by metacognitive evaluation activities (towards the end of the learning sessions); tiny fraction of metacognitive evaluation activities towards the end of the sessions
* Pattern/strategy 2: focus on summative assessment with indicators of trial-and-error learning (large percentage of incorrectly solved exercises)
* Pattern/strategy 3: focus on reading lecture materials with tiny fraction of formative assessment
* Pattern/strategy 4: focus on the course videos, with not negligible amount of formative assessment activities; small fraction of metacognitive evaluation activities at the beginning of the learning sessions

Student clusters, identified in Step 2:

* Cluster 1 (22, 7.59%): the most active group, practicing a variety of learning strategies, with strategies 3 and 2 being the most prominent; they have the highest final exam score
* Cluster 2 (70, 24.13%): similar to cluster 1, but with lower activity level and a reverse importance of strategies 2 and 3 - this group puts somewhat more emphasis on summative assessment (strategy 2), than on reading lecture materials (strategy 3), while in cluster 1 the opposite is the case; furthermore, compared to cluster 1, it is significantly less practicing strategy 1; this group has somewhat lower final exam and slightly higher midterm scores than group 1, but the difference is not statistically significant
* Cluster 3 (133, 45.86%): this is the largest group, with strategy 2 as the most dominant one, though they experimented with other learning strategies, as well; their overall level of activity and exam scores are further below the previous two groups (clusters 1 and 2)
* Cluster 4 (65, 22.41%): the least active group, practicing almost exclusively strategy 2, and having the lowest scores both on the midterm and the final exam.

The table below allows for a more detailed comparison of the student clusters.

================ ==================== ================== ================ ================

attributes Median; (Q1, Q3) Median; (Q1, Q3) Median; (Q1, Q3) Median; (Q1, Q3)

================ ==================== ================== ================ ================

cluster 1 2 3 4

n.students 22 70 133 65

seq.clust1 12.5; (11, 15) 6.5; (4, 10) 3; (1, 5) 1; (0, 3)

seq.clust2 26; (23.25, 29) 21; (18, 23) 18; (16, 21) 10; (7, 13)

seq.clust3 29.5; (24, 33.5) 18; (15, 21) 8; (5, 11) 3; (2, 5)

seq.clust4 12; (9, 13.75) 9; (7, 12) 4; (2, 7) 1; (0, 2)

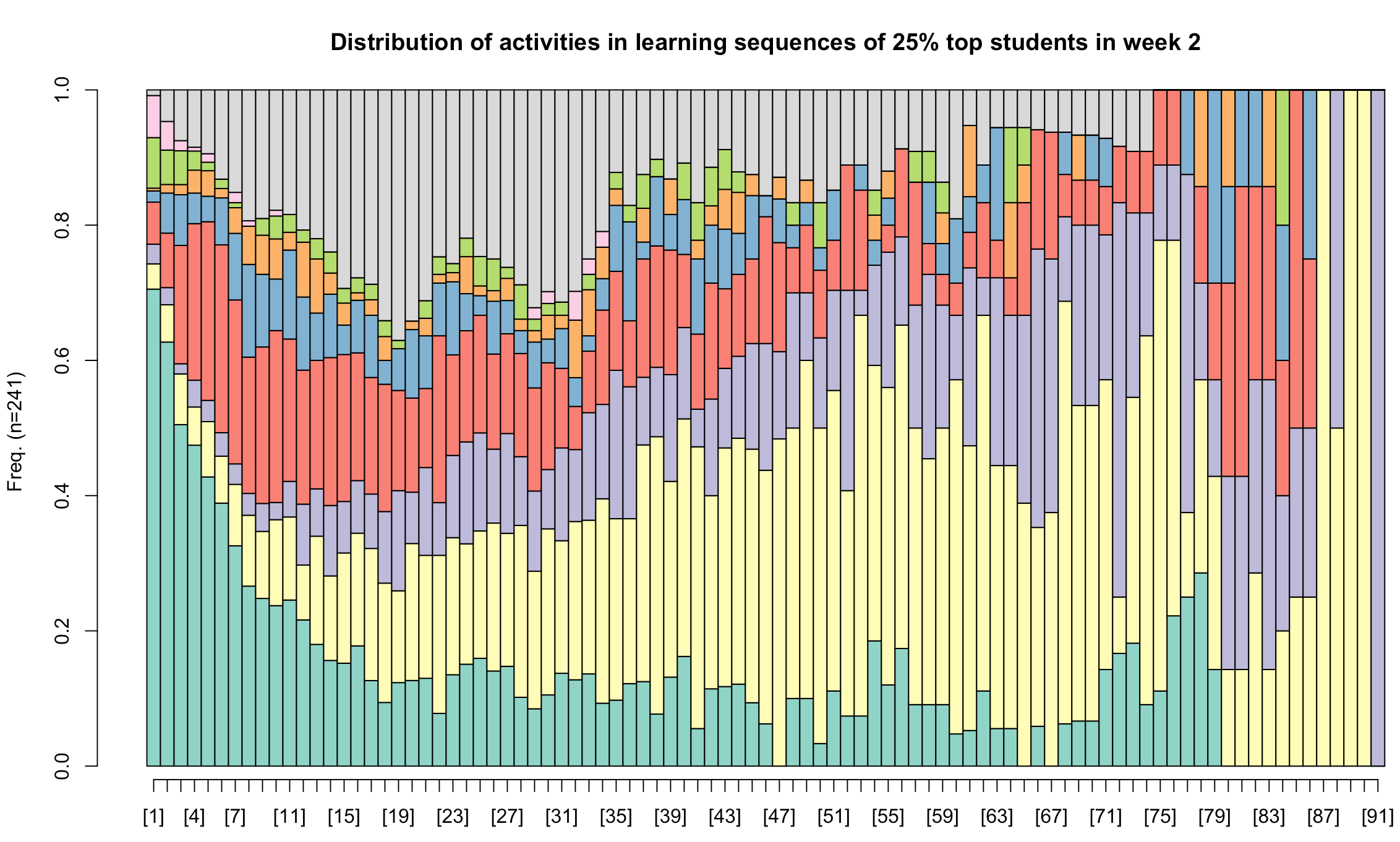
seq.total 78.5; (74.25, 89) 55; (50, 59) 34; (29, 39) 17; (13, 20)

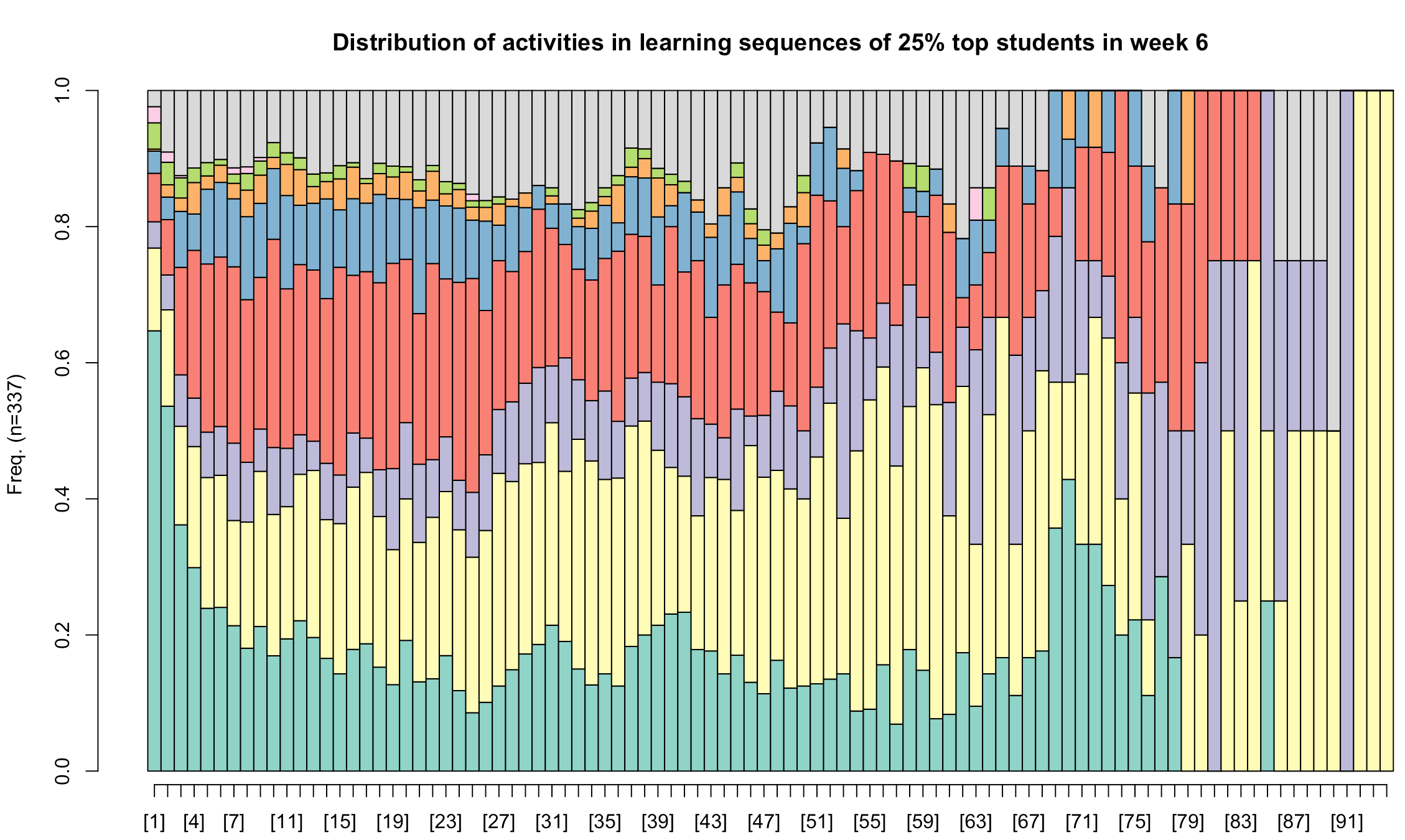
midterm.score 15; (13, 17) 16; (13, 17) 13; (11, 16) 11; (8, 15)

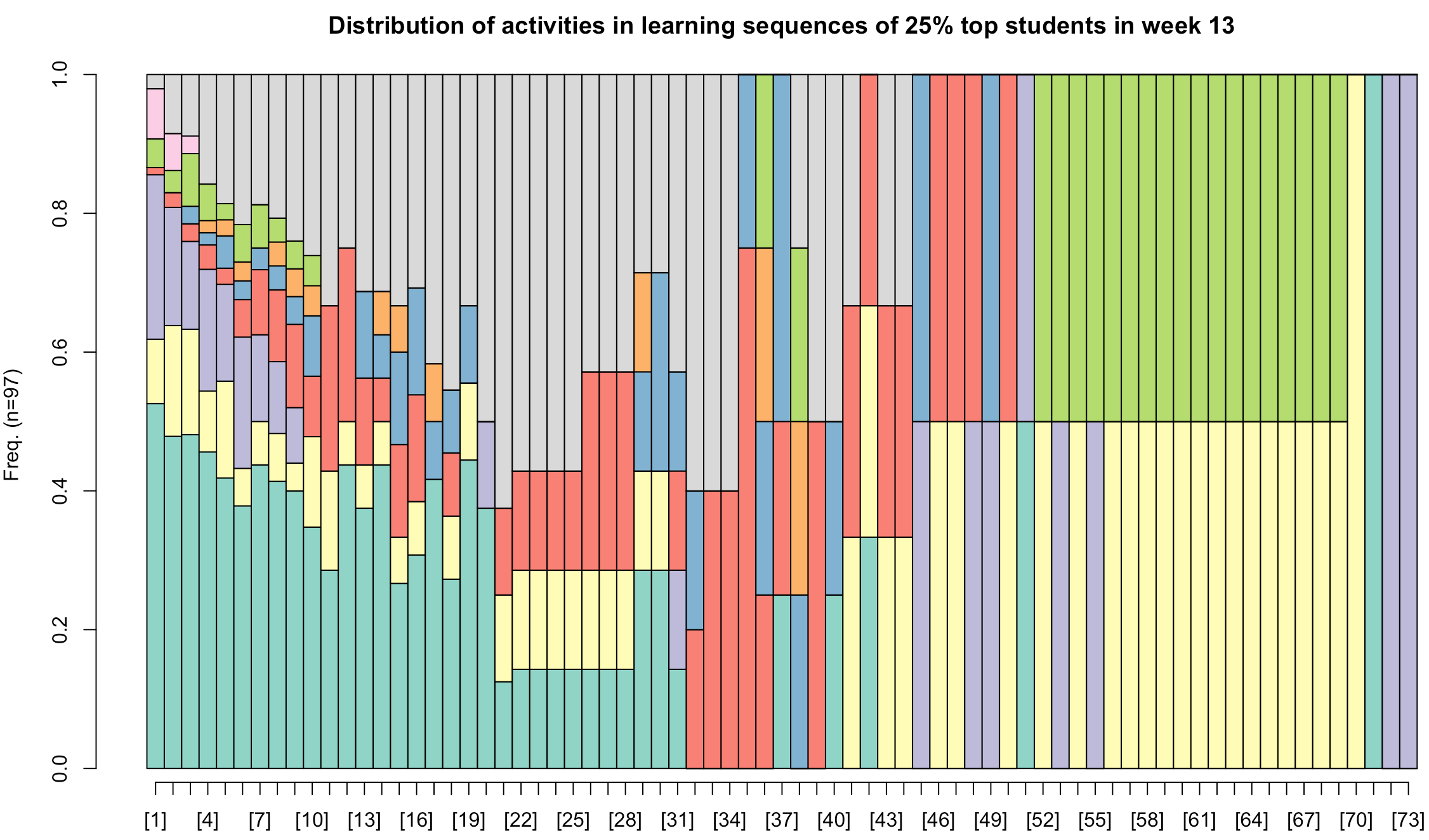
final.exam.score 22.5; (17.25, 29.75) 21; (15.25, 30.75) 16; (13, 24) 14; (10, 18)

================ ==================== ================== ================ ================

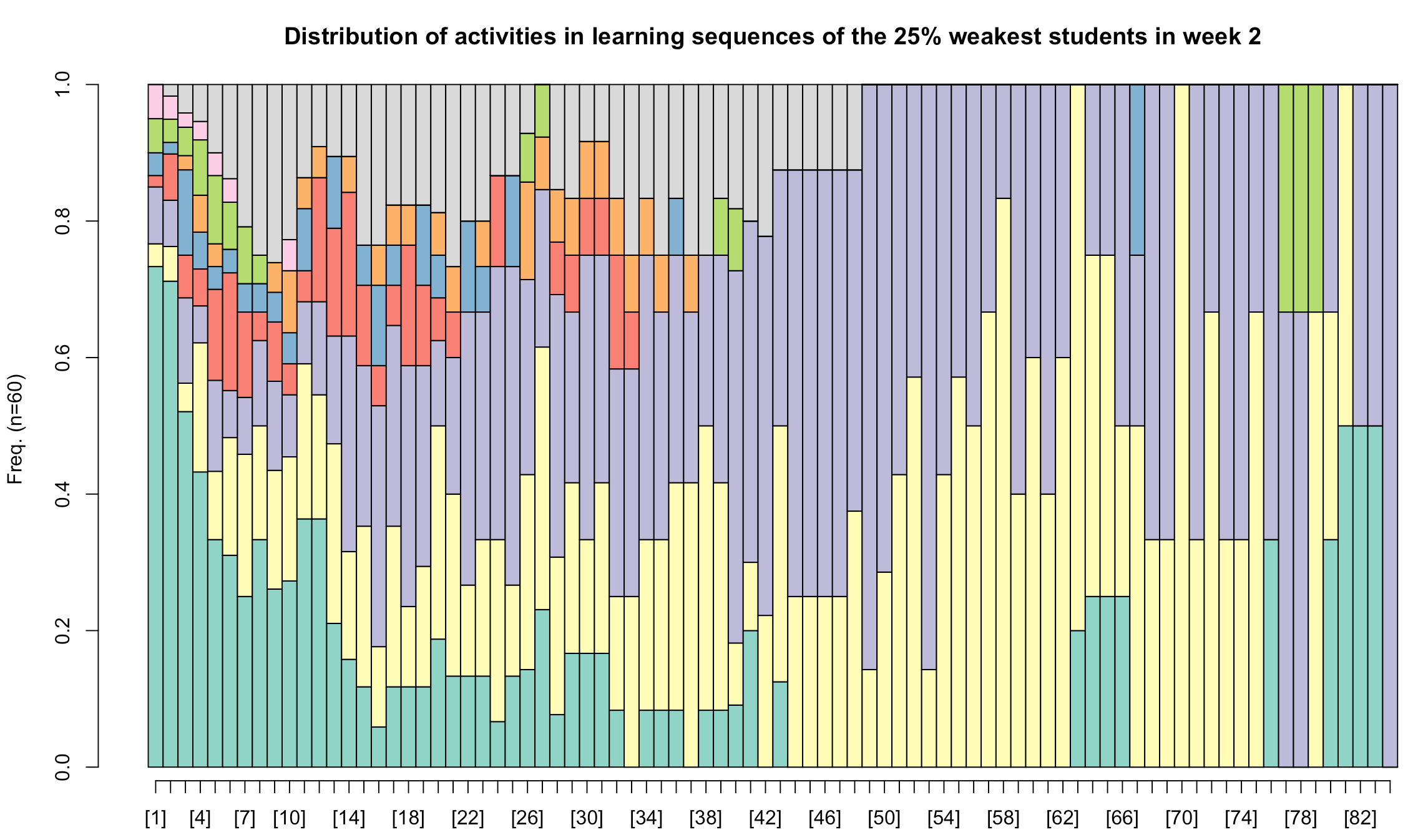
**Sequence graphs for the 1st, 6th and 13th week, for 25% top students**







**Sequence graphs for the 1st, 6th and 13th week, for 25% weakest students**

****

