

1. پژوهشگری قصد دارد دو توالی آمینو اسیدی زیر را از طریق الگوریتم Needleman–Wunsch، Align کند. او جدول زیر را کشیده است و می خواسته از ماتریس PAM250 استفاده کند. جدول را کامل کرده و هم ردیفی یا هم ردیفی های بهینه را بنویسید.
(15 نمره مثبت برای جدول درست و 15 نمره برای همردیفی درست)

توالی 1: AWNNTCS

توالی 2: WNSTSS

Gap Open: 1

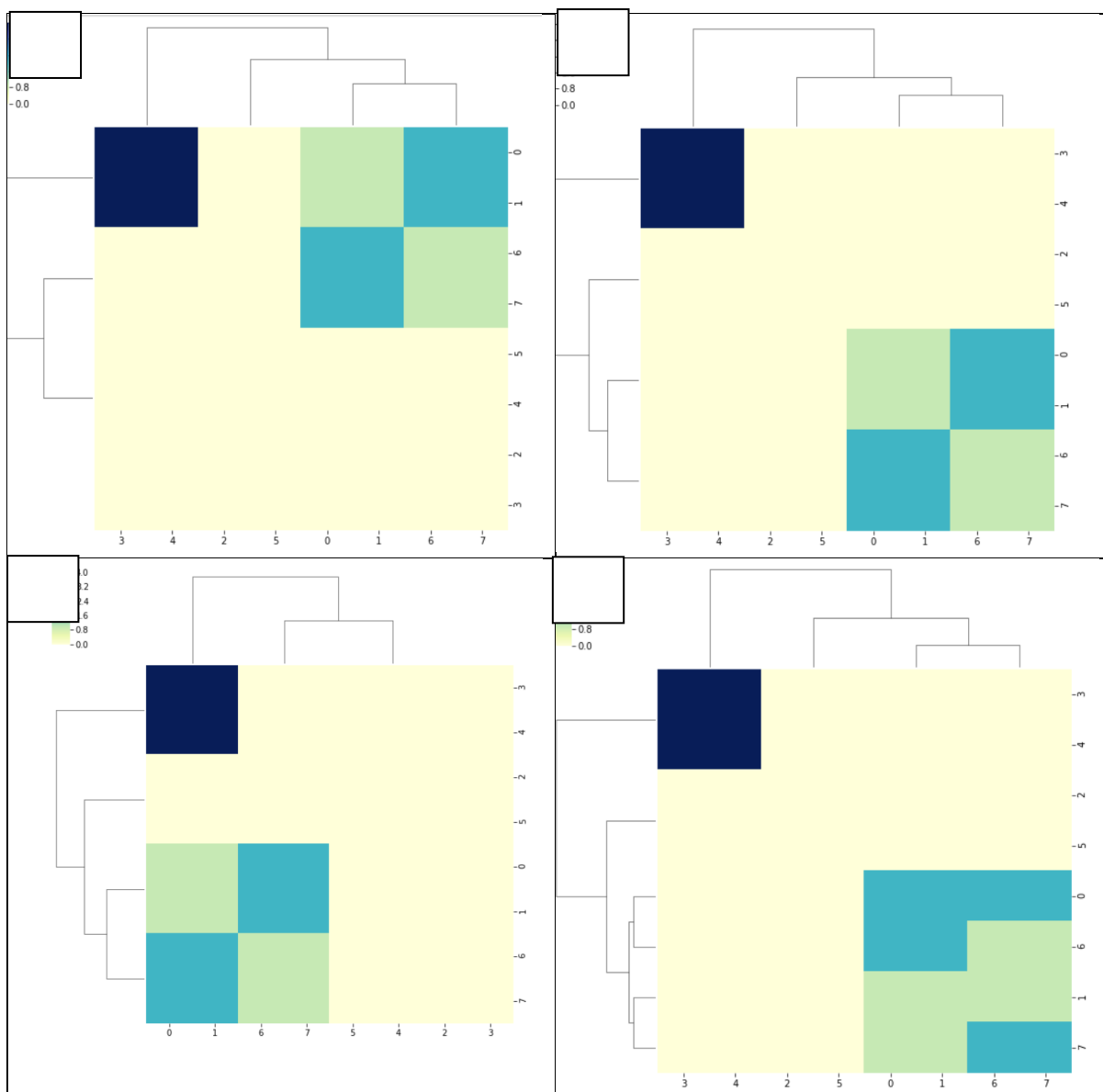
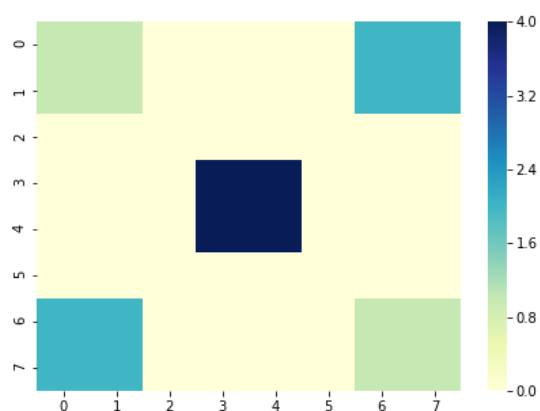
Gap Extend: 0.5

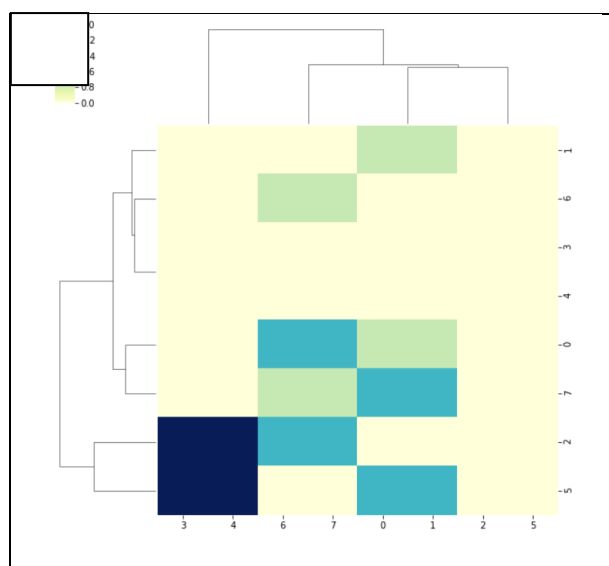
End Gap Open: 0.25

End Gap Extend: 0.25

		A	W	N	N	T	C	S
	O							
W								
N								
S								
T								
S								
S								

2. در نقشه حرارتی (HeatMap) زیر حالت بدون اعمال dendrogram یک آرایه دو بعدی را مشاهده می کنید. حال پس از اعمال dendrogram کدام یک از نمودار ها ایجاد خواهند شد ؟
(10 نمره مثبت، 3 نمره منفی)





3. جدول زیر تعداد خوانش ژن های 1 تا 4 را در گونه های A تا G در یک بررسی RNAseq نشان می دهد.

Read Count	G1	G2	G3	G4
A	0	14	25	19
B	4	4	13	24
C	16	5	2	21
D	3	7	14	17
E	1	2	1	27
F	3	1	9	20
G	0	3	2	30

از این دیتا برای Hierarchical Clustering این گونه ها استفاده کنید. 50 نمره

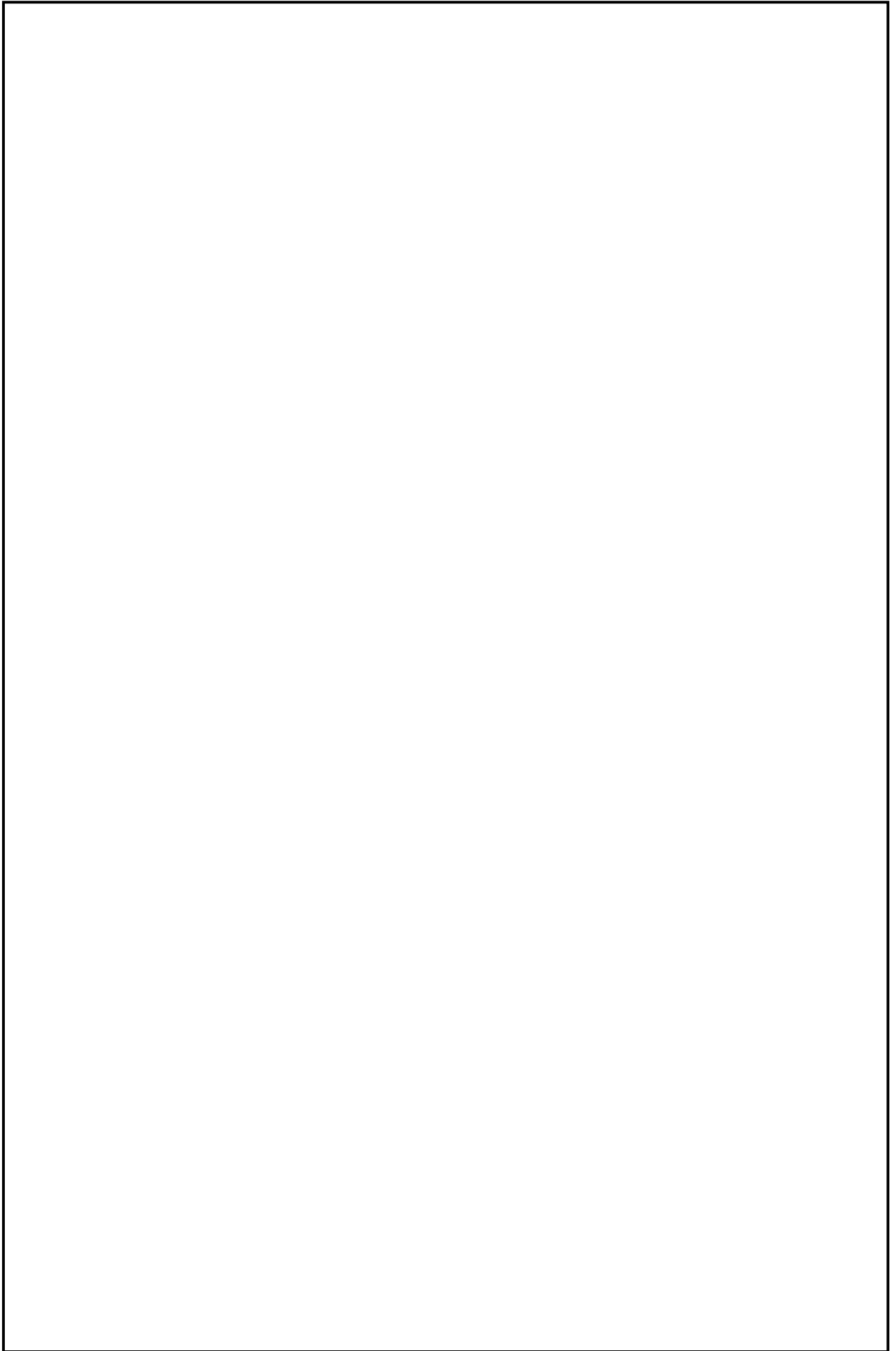
Clustering Method: Single-linkage

Distance Measurement Method: Euclidean

جدول اولین Distance Matrix برای شما رسم شده است.
در صورت نیاز به بروزرسانی، هر مرحله را در پاسخ خود یادداشت کنید (در صورت نیاز جدول های مراحل بعدی را رسم و تکمیل کنید).

Dendrogram نهایی را به همراه فاصله ها، رسم کنید.

Distance	A	B	C	D	E	F	G
A							
B							
C							
D							
E							
F							
G							



4. با توجه به ارتباط PDF (تابع چگالی احتمال) و CDF (تابع تجمعی احتمال) صحت گزاره های زیر را بررسی کنید. (هر مورد 5 نمره و نمره منفی برابر)

	الف) شکل CDF تابع چگالی احتمال مرتبط با توزیع نرمال سیگموئید خواهد بود.
	ب) با افزایش انحراف معیار توزیع، حداکثر مقدار شیب نمودار CDF افزایش خواهد یافت.
	ج) مشتق دوم تابع PDF در مقدار امید ریاضی توزیع پواسون مثبت خواهد بود.
	د) در بازه اعداد حقیقی، مشتق تابع PDF حداقل یکبار از محور X ها عبور خواهد کرد.
	ه) مقدار مطلق تابع CDF می تواند بیشتر از عدد یک شود.