Coh	ort 3	1 - Group 4						
√an	essa	Gleason, Alistair	Marsden, Oliv	ier Rochaix, Edua	ırd Stalmakov			
Сар	stone	Project						
Pro	ject N	lanagement Plan						
Χ	ID bi	oad topic - ecomm	erce > music s	streaming service				
ΙP	Iden	tify 3 data sets (plu	s the 1 census	dataset)				
		all teammates exp	lore and find a	potential concept				
Χ	crea	te GitHub repositor	•					
		https://github.com/	<u>EduardStalma</u>	kov/Capstone-Pro	<u>ect</u>			
IP	com	come up with 5-10 questions about data in your industry						
		stored on GitHub a						
		one question has t	· ·	ons, classifications	, or find clusters	of some		
		formalize question			1 50/40			
		we'll go through th	e questions a	few times during w	eek of 9/19			
				D 0/00/0	_			
	Crea	<mark>ite elevator pitch fo</mark>	rmal summary	to sen Due 9/20/2	2			
				D 0/00/0	0			
IP	crea	t <mark>e project manager</mark>	•	Due 9/23/2				
	stored on GitHub as "ProjectManagementPlan.pdf"							
	Cotti	ng up plotform for i	araja at	Drainated (<mark>//10_0</mark> _22			
		ng up platform for _l		Projected 9	// 19-9- 23			
	Reg	Regarding Kafka and Cloud ETL:						
		must have Kafka Producer Stream						
		must configure and consume topics capture the kafka stream in a databrick						
					ta laka			
		use the databrick tuse azure data fac	•			~ SOI		
		Specifically to CI	•		uala lane lu azul	e oul		
				ce diagram of data	ninelines			
			services are	•	pipelliles			
			hings are getti	-				
		identity flow t	imiga are getti	ing moved				
	crea	te napkin drawings	and feedback	for the visualization	ns			
	orca	save all screensho						
		stored on GitHub a						
		2.0.00 011 0111 100 0	Tioddiizatic					
	crea	te napkin drawings	and feedback	for the dashboard				
	5.00	save all screensho			nt			
		stored on GitHub a		and an a docume				

reate repeatable ETL report							
include Kafka							
cloud ETL							
any other ETL you deem necessary							
stored on GitHub as "RepeatableETLReport.pdf"							
reate dashboard							
can be in Dash or Power BI w Python							
include the Dash dashboard or Power BI Report files in a folder in your repo called "dashboard"							
there is any code involved for the deliverables:							
Make a folder called "code"							
Store all of your code in this code folder							
Keep in mind that code must be kept clean and readable							
reate Project Technical Report							
Save this as ProjectTechnicalReport.pdf							
resentation Slides							
you can use whatever presentation platform you prefer.							
export your slides to PDF for review.							
stored on GitHub as "CapstonePresentationSlides.pdf"							
hould include the following tools and concepts							
ython							
andas or pyspark							
ML concepts							
afka							
eliverables							
- Control - Cont							
fka Messaging							
ata Factory Pipeline							
ata Lake Usage							
PI Use							
chine Learning Model creation							
ashboard (That incorporates some Machine Learning predictions) - can be Dash or Power BI w Python							
/ritten Technical Report							
ninute group presentation							
minute group presentation							

GitHub repository must include:						
5-10	ques	stions about data				