

Q1.

I have read the syllabus with the grading policy as well as the academic integrity form for this online quarter and will abide by the rules and complete my coursework with integrity.

Q2.

Hello Bioinformatics

Q3.

>gi|6978799|ref|NP\_036683.1| early growth response 1; nerve growth factor-induced gene [Rattus norvegicus] 508  
>gi|45768856|gb|AAH67618.1| Serum/glucocorticoid regulated kinase [Danio rerio] 433  
>gi|45768786|gb|AAH68134.1| Unknown (protein for MGC:95907) [Mus musculus] 423  
>gi|27923854|sp|P59241|STK6\_RAT Serine/threonine kinase 6 (Aurora-A) (ratAurA) 397  
>gi|45768720|gb|AAH67812.1| Cyclin L1 [Homo sapiens] 526  
>gi|45768758|gb|AAH68160.1| Cdk7 protein [Mus musculus] 346  
>gi|45219906|gb|AAH66834.1| Mastl protein [Mus musculus] 671  
>gi|18202599|sp|Q63796|M3KC\_RAT Mitogen-activated protein kinase kinase kinase 12 (MAPK-upstream kinase) (MUK) 888  
>gi|4835224|emb|CAB42902.1| protein kinase ATN1 like protein [Arabidopsis thaliana] 370  
>gi|40787731|gb|AAH64804.1| SLK protein [Homo sapiens] 617  
>gi|18202068|sp|O55173|PDPK\_RAT 3-phosphoinositide dependent protein kinase-1 (Protein kinase B kinase) (Pb kinase) 559  
>gi|34191428|gb|AAH36504.2| C9orf96 protein [Homo sapiens] 700  
>gi|29747774|gb|AAH50806.1| Gene model 711, (NCBI) [Mus musculus] 587  
>gi|28856169|gb|AAH48033.1| Serine/threonine kinase 3 (STE20 homolog, yeast) [Danio rerio] 492  
>gi|20071571|gb|AAH26466.1| Unknown (protein for IMAGE:4485517) [Mus musculus] 202  
>gi|45709347|gb|AAH67695.1| Unknown (protein for MGC:85918) [Danio rerio] 320

Q4.

>gi|6978799|ref|NP\_036683.1| early growth response 1; nerve growth factor-induced gene [Rattus norvegicus]

MDNYPKLEEMMLLSNGAPQFLGAAGTPEGSGGNNSSSSSSSSSSSGGGGGGGSNSGSSAFNP

QGEPSEQPYEHLTTESFSDIALNNEKALVETSYPSQTTRLPPITYTGRFSLEPAPNSGNT

LWPEPLFSLVSGLVSMNTPPTSSSSAPSPAASSSSSSASQSPPLSCAVPSNDSSPIYSAAP

TFPTPNTDIFPEPQSQAFFGSAGTALQYPPPAYPATKGGFQVPMIPDYLFPPQQQGDLSLG

TPDQKPFQGLNRTQQPSLTPLSTIKAFATQSGSQDLKALNNTYQSQLIKPSRMRKYPNR

PSKTPPHERPYACPVESCDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLTTHI

RTHTGEKPFACDICGRKFARSDERKRHTKIHLRQKDKKADKSVVASSAASSLSSYPSPVA

TSYPSPATTSFPSPVPTSYSSPGSSSTYPSPAHSGFPSPSVATTYASVPPAFPAQVSTFQS

AGVSNSFSTSTGLSDMTATFSPRTIEIC

>gi|45768786|gb|AAH68134.1| Unknown (protein for MGC:95907) [Mus musculus]

MSTRNCQGTDSVIKHLDTIPEDKKVRVQRTQSTFDPFEKPANQVKRVHSENNACINFKSS  
SAGKESPKVRRHSSPSSPTSPKFGKADSYEKLEKLGEFSYATVYKGKSKVNGKLVALKVI  
RLQEEEGTPFTAIREASLLKGLKHANIVLLHDIHTKETLTLVFEYVHTDLCQYMDKHPG  
GLHPDNVKLFLFQLLRGLSYIHQRYILHRDLKPQNLLISDTGELKLADFGLARAKSVPSH  
TYSNEVVTLWYRPPDVLLGSTYSTCLDMWGVGCIFVEMIQGVAAFPGMKDIQDQLERIF  
LVLGTPNEDTWPGVHSLPHFKPERFTVYSSKSLRQAWNKL SYVNHAEDLASKLLQCSPKN  
RLSAQAALSHEYFSDLPPRLWELTDMSSIFTVPNVRLQPEAGESMRAFGKNNSYGKSLSN  
SKH

>gi|27923854|sp|P59241|STK6\_RAT Serine/threonine kinase 6 (Aurora-A) (ratAurA)

MDRCKENCVSRPVKSTVPFGPKRVLVTEQIPSQHPGSASSGQAQRVLCPSNSQRVPPQAQ  
KPVAGQKPVLKQLPAASGPRPASRLSNPQKSEQPQPAASGNNSEKEQTSIQKTEDSKKRQ  
WTLEDFDIGRPLGKGKFGNVYLAREKQSKFILALKVLFKVQLEKAGVEHQLRREVEIQSH  
LRHPNILRLYGYFHDATRVYLLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYC  
HSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTLCGTLDYQPPEMIEGRMHD  
EKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDLISRLKHN  
SSQRLTLAEVLEHPWIKANSSKPPTGHNSKEATSKSS

>gi|45768758|gb|AAH68160.1| Cdk7 protein [Mus musculus]

MAVDVKSRAKRYEKLDLFLGEGQFATVYKARDKNTNQIVAIKKIKLGHRSEAKDGINRTAL  
REIKLLQELSHPNIIIGLLDAFGHKSNISLVFDFMETDLEVIKDNSLVLTPSHIKAYMLM  
TLQGLEYLHQHWILHRDLKPNNLLLDENGVLKLADFGLAKSFGSPNRAYTHQVVTRWYRA  
PELLFGARMYGVGVDMWAVGCILAEALLRVLPFLPGDSLDLQTRIFETLGTPTTEEQWPDM  
CSLPDYVTFKSFPGVPLQHIFIAAGDDLLELIQGLFLNPCTRRTTASQALKTKYFSNRPG  
PTPGCQLPRPNCPVEALKEPANPTVATKRKRAEAEQGILPKKLIF

>gi|45219906|gb|AAH66834.1| Mastl protein [Mus musculus]

SMSKPKQDYSRTPGQVLSLISSLGFFTPVGEKDQDSANMFSA PKSAAQLSRGFICPMSVD  
QKEPTSYSSKLLKSCFETLSSNPEIPVKCLTSNLLQCRKRLGTSSSTSSQSHTFVSSVESE

CHSNPKWERDCQSTESSGCAMSWNAVEMLYAKSTSAIKTKTELELALSPIHDSSAIPAAG  
SNQVTLPRKCFREISWEARDPDNENMTIDKGQSGFCQSSQRSVNSSATSEEHLGKRNYKR  
NFHLVDSSPCQEIMQSKKNCTEYEANKERQGCRANQSTGLTTEVQNLKLSGCESQQLDYA  
NKENIVTYLTDRQTPEKLHIPTIAKNLMSELDEDRELSSKKDCLSSNSVCSEDEDRAKTT  
CVDSDSFFPGVSMMESSLEIQALEPDKSIRDYSFEENPTEDLFVLPKCQENSLPQDDCHA  
CIQDSSQVSAHPSKAPKALTSKINVVAFRSFNHINASTNSEPSKISITSLDAMDIDSYDY  
SGSYPMASVSPTEKGRHYTSHQTPNQVKLGTSYRTPKSVRRGAAPVDDGRILGTPDYLAPE  
LLLGTAHGPAVDWWALGVCLFEFLTGIPIFNDETQPVFQNILKRDIPWPEGEEKLSDNA  
QSAMDMLLTIDDSKRAGMRELKQHPLFSEVDWENLQHQTMPFVPQPDDDETDTSYFEARNN  
AQHLTISGFSL

>gi|18202599|sp|Q63796|M3KC\_RAT Mitogen-activated protein kinase kinase kinase 12 (MAPK-upstream kinase)  
MUK)

MACLHETRTPSPSFGGFVSTLSEASMRKLDPDTSCTPEKDLTPTQCVLRDVVPLGGQGG  
GGPSPSPGGEPPEPFANSVLQLHEQDTGGPGGATGSPESRASRVRADEVRLQCQSGSGF  
LEGLFGCLRPVWTMIGKAYSTEHKQQQEDLWEVPFEEILDQWVGSGAQQGAVFLGRFHGE  
EVAVKKVRDLKETDIKHLRKLKHPNIITFKGVCTQAPCYCILMEFCAQQQLYEVLRAGRP  
VTPSLLVDWSMGIAGGMNYLHLHKIIHRDLKSPNMLITYDDVVKISDFGTSKELSDKSTK  
MSFAGTVAWMAPEVIRNEPVSEKVDIWSFGVVLWELLTGEIPYKDVDSSAIIWGVGSNSL  
HLPVPSSCPDGFKILLRQCWNRKPRNRPSFRQILLHLDIASADVLSTPQETYFKSQAQEW  
EEVKLHFEEKIKSEGTCLHRLEEELVMRRREELRHALDIREHYERKLERANNLYMELNALM  
LQLELKERELLRREQALERRCPGLLKSHTSRSLHGNTEKLIKRNVPQKLSPHSKRPD  
ILKTESLLPKLDAALSGVGLPGCPKAPPSPGRSRRGKTRHRKASAKGSCGDLPLRAALP  
PHEPGGLGSPGGLGVGPTAWDASPPALRGLHHDLLLKMSSSSPDLLSAALGARGRGATG  
GARDPGSPPPPQGDTPPSEGSAAGSTSPDSPGGAKGEPPPPVGPGEVGLLGTGREGTTG  
RGGSRAGYQHLTPAALLYRAAVTRSQKRGISSEEEEEGEVDSEVELPPSQRWPQGPNMRQS  
LSTFSENPSDVEEGTASEPSPSGTPEVGSTNTDERPDERSDDMC SQGSEIPLDLPTSEV

VPERETSSLPMQHQQDDQGNPEDSDCDSTELDNSNSIDALPPPASLPP

>gi|18202068|sp|O55173|PDPK\_RAT 3-phosphoinositide dependent protein kinase-1 (Protein kinase B kinase) (Pk kinase)

MARTTSQLYDAVPIQSSVVLCSPPSMVRSQTEPSSSPGIPSGVSRQGSTMDGTTAEAR  
PSTNPLQQHPAQLPPQPRKKRPEDFKFGKILGEGSFSTVVLARELATSREYAIKILEKRH  
IIKENKVPYVTRERDVMSRLDHPFFVKLYFTFQDDEKLYFGLSYAKNGELLKYIRKIGSF  
DETCTRFYTAEIVSALEYLHGKGIIHRDLKPENILLNEDMHIQITDFGTAKVLSPDSKQA  
RANSFVGTAQYVSPELLTEKSACKSSDLWALGCIYQLVAGLPPFRAGNEYLIFQKIIKL  
EYDFPEKFFPKARDLVEKLLVLDA TKRLGCEEMEGYGPLKAHPFFESITWENLHQQTTPK  
LTAYLPAMSEDDDCYGNYDNLLSQFGCMQVSSSSSSSHSLCAVDASLPQRSGSNIEQYIH  
DLDTNSFELDLQFSEDEKRLLEKQAGGNPWHQFVENNLILKMGPVDKRKGLFARRRQLL  
LTEGPHLYYVDPVNKVLKGEIPWSQELRPEAKNFKTFFVHTPNRTYYLMDPSGNAHKWCR  
KIQEVWRQQYQSSPDAAVQ

>gi|29747774|gb|AAH50806.1| Gene model 711, (NCBI) [Mus musculus]

MDYYSQGTFQNIMENKRKLKAVVDTEWMHTMLSQVLDAIEYLHKLNIVHRNLKPSNIVLV  
NSGYCKLQDMSSQALMTHEAKWNVRAEEDPCQKSWMapealkfsfstksdiwslgciild  
MATCSFLNDTEAMQLRKAI RHHPGSLKPILKTMEEKQIPGTDVYYLLLPFMLHINPSDRL  
AIKDVMQVTFMSNSFKSSSV ALNMQRQKVPIFITDVLLEG NMANILDVMQNFSSRPEVQL  
RAINKLLTMPEDQLGLPWPTELLEVISIIKHGRILDILLSTCSLLLRVLGQALAKDPE  
AEIPRSSLIISFLMDTLRSHPNSERLVNVVYNVLAISSQGQISEEEEEGLFQLAQENL  
EHFQEDRDICLSILSLLWSLLVDVVTVDKEPLEQLSGMVTWVLATHPEDVEIAEAGCAVL  
WLLSLLGCIKESQFEQVVVLLLR SIQLCPGRVLLVNNAFRGLASLAKVSELVAFRIVVLE  
EGSSGLHLIQDIYKLYKDDPEVVENLCMLLAHLTSYKEILPEMESGGIKDLVQVIRGRFT  
SSLELISYADEILQVLEANAQPGLQEDQLEPPAGQEAPLQGEPLFRP

>gi|20071571|gb|AAH26466.1| Unknown (protein for IMAGE:4485517) [Mus musculus]

PTRPTRLIVSNFSQAKQKSHLVDPQILRDQSRLAPEIITATQYKKCDEFQTGILYEMLH  
LPNPFDENPELKEKEYTRTDLPRIPLRSPYSWGLQQLASCLLNPNPSEILISDAKGILQ

CLLWGPREDLFQIFTTSATLAQKNALLQNWLDIKRTLLMIKFAEKSLDREGGISLEDWLC

AQYLAFATTDLSLSYIVKILQYR

Q5a.

MDNYPKLEEMMLLSNGAPQFLGAAGTPEGSGGNNSSSSSSSSSSSGGGGGGGGSNSGSSAFNPQGEPSEQPYEH  
TTESFSDIALNNEKALVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPLFSLVSGLVSMTNPPTSSSSA  
SPAASSSSSSASQSPPLSCAVPSNDSSPIYSAAPTFTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPATKGGFQ  
PMIPDYLFPQQQGDLSLGTDPDQKPFQGLENRTQQPSLTPLSTIKAFATQSGSQDLKALNNTYQSOLIKPSRM  
KYPNRPSKTPPHERPYACPVESCDRRFSRDELTRHIRIHTGQKPFQCRICMRNFSRSDHLTTHIRHTGEKP  
ACDICGRKFARSDEKRHTKIHLRQKDKKADKSVVASSAASSLSSYSPVATSYPSPATTSFSPSPVPTSYSPP  
SSTYPSPAHSGFSPSPSVATTYASVPPAFPAQVSTFQSAGVSNSFSTSTGLSDMTATFSPRTIEIC@MTIQTETS  
SAPDLTYSKTRGLVANLSAFMKQRKMGLNDFIQKLSANSYACKHPEVQSILNLTPPDVELMNSNPSPPPS  
SQQINLGPSSNPTAKPSDFDFLKVIGKGSFGKVLLARHRSDEKFYAVKVLQKKAILKKKEEKHIMSENVLL  
NVKHPFLVGLHYSFQTTDKLYFVLDYINGGELFYHLQRERCFLERARFYAAEIASALGYLHSLNIVYRDL  
PENILLDSQGHILTDGFLCKENIEPNGTTSTFCGTPEYLAPEVLHKQPYDRTVDWWCLGAVLYEMLYGLPP  
YSRNTAEMYDNILNKPLQLKPNISNAARHLLEGLLQKDRTKRLGFTDDFTEIKNHMFFSPINWDDLNAKKL  
PPFNPNTGPNDLRHFDPFTDEPVNSIGCSPDSALVTSSITEATEAFLGFSYAPAMDSYL@MSTRNCQGT  
SVIKHLDTIPEDKKVRVQRTQSTFDPFEKPANQVKRVHSENNACINFKSSSAGKESPKVRRHSSPSSPTSPKF  
KADSYEKLEKLGEFSYATVYKKGSKVNGKLVALKVIRLQEEEGTPFTAIREASLLKGLKHANIVLLHDIHT  
ETLTLVFEYVHTDLCQYMDKHPGGLHPDNLKFLFLQLLRGLSYIHQRYILHRDLKPQNLLISDTGELKLAD  
GLARAKSVPSHTYSNEVVTLWYRPPDVLLGSTYEYSTCLDMWGVGCIFVEMIQGVAAFPGMKDIQDQLERI  
LVLGTPNEDTWPGVHSLPHFKPERFTVYSSKSLRQAWNKL SYVNHAEDLASKLLQCSPKNRLSAQAALSH  
YFSDLPPRLWELTDMSSIFTVPNVRLQPEAGESMRAFGKNNSYGKSLSNSKH@MDRCKENCVSRPVKSTV  
FGPKRVLVTEQIPSQHPGSASSGQAQRVLCPSNSQRVPPQAQKPVAGQKPVLKQLPAASGPRPASRLSNPQ  
SEQPQAASGNSEKEQTSIQKTEDSKKRQWTLEDFDIGRPLGKGKFGNVYLAREKQSKFILALKVLFKVQ  
EKAGVEHQLRREVEIQSHLRHPNLRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELA  
ALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTLCGTLDYQPPEMIEGRMHDEKVDLW  
LGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDLISRLKHNSSQRLTLAEVLEHPWIKAN  
SKPPTGHNSKEATSKSS@MASGPHSTATAAAAASSAAPSAGGSSSGTTTTTTTTTGGILIGDRLYSEVSLTID  
SLIPEERLSPTPSMQDGLDLPSETDLRILGCELIQAAGILLRLPQVAMATGQVLFHRFFYSKSFVKHSFEIVA  
ACINLASKIEEAPRRIRDLINVFHHLRQLRGKRTSPPLILDQNYINTKNQVIKAERRVLKELGFCVHVKHPHK  
IVMYLQVLECERNQTLVQTAWNYMNDLRTNVFVRFPETIACACIYLAARALQIPLPTRPHWFLFLGTTE  
EIQEICIETRLRYTRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKA  
KSPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSTRSRSRSHTPRRHYNRRSRSGTY  
SRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRRKRSRSQSKSRDHSDAAKKHRHERG  
HRDRRERSRSFERSHKSKHHGGSRSRSGHGRHRR@MAVDVKSRAKRYEKLDLFLGEGQFATVYKARDKNTN  
IVAIIKKIKLGHRSEAKDGINRTALREIKLLQELSHPNIIIGLLDAFGHKSNI SLVDFMETDLEVIKDNSLVLT  
HIKAYMLMTLQGLEYLHQHWILHRDLKPNNLLLDENGVLKLADFLAKSFGSPNRAYTHQVVTRWYRAP  
LLFGARMYGVGVDMWAVGCILAELLRVPFLPGDSLDLQLTRIFETLGTPTTEEQWPDMCSLPDYVTFKSFP  
VPLQHIFIAAGDDLLELIQGLFLFNPCRTTASQALKTKYFSNRPGPTPGCQLPRPNCPVEALKEPANPTVAT  
RKRAEAELEQGILPKKLIF@SMSKPKQDYSRTPGQVLSLISSLGFFTPVGEKDQDSANMFSAPKSAAQLSRGF  
CPMSVDQKEPTSYSSKLLKSCFETLSSNPEIPVKCLTSNLLQCRKRLGTSSTSSQSHTFVSSVESECHSNPKW  
RDCQSTESSGCAMSWNAVEMLYAKSTSAIKTKTELELALSPIHDSSAIPAAGSNQVTLPRKCFREISWEARD  
DNENMTIDKGQSGFCQSSQRSVNSSATSEEHLGKRNYKRNHFLVDSSPCQEIMQSKKNCTEYEANKERQG  
RANQSTGLTTEVQNLKLSGCESQQLDYANKENIVTYLTDRQTPEKLHIPTIAKNLMSELDEDRELSSKKDCL  
SNSVCSDERALKTTCDVSDSSFPVSMMESSLEIQALEPDKSIRDYSFEENPTEDLFLVLPKCQENSLPQDDC  
ACIQDSSQVSAHPSKAPKALTSKINVVAFRSFNSHINASTNSEPSKISITSLDAMDISYDYSGSYPMAVSPTEK  
RHYTSHQTPNQVKLGTSYRTPKSVRRGAAPVDDGRILGTPDYLAPELLGTAHGPAVDWWALGVCLFEFL

GIPIFNDETPQQVFQNILKRDIPWPEGEEKLSDNAQSAMDMLLTIDDSKRAGMRELKQHPLFSEVDWENLQ  
QTMPFVPQPDDDETDTSYFEARNNAQHLTISGFSL@MACLHETRTPSPSFGGFVSTLSEASMRKLDPDTSDCT  
EKDLTPTQCVL RDVVPLGGQGGGGPSPSPGGEPPPEPFANSVLQLHEQDTGGPGGATGSPESRASRVRADE  
RLQCQSGSGFLEGLFGCLR PVWTMIGKAYSTEHKQQQEDLWEVPFEEILDLQWVGSGAQGA VFLGRFHGE  
VAVKKVRDLKETDIKHLRKLKHPNIITFKGVCTQAPCYCILMEFCAQGQLYEVL RAGRPVTPSLLVDWSM  
IAGGMNYLHLHKIIHRDLKSPNMLITYDDVVKISDFGT SKELSDKSTKMSFAGTVAWMAPEVIRNEPVSEK  
DIWSFGVVWLWELLTGEIPYKDVDSSAIIWGVGSNSLHLPVPSSCPDGFKILLRQCWNRKPRNRP SFRQILLHL  
IASADVLSTPQETYFKSQA EWREEVKLHFEEKIKSEGTC LHRLEEELVMRRREELRHALDIREHYERKLERAN  
LYMELNALMLQLELKERELLRREQALERRCPGLLKSHTSRSL LHGNTMEKLIKRNVPQKLSPHSKRPDIL  
TESLLPKLDAALSGVGLPGCPKAPPSPGRSRRGKTRHRKASAKGSCGDL PGLRAALPPHEPGGLGSPGGLG  
GPTAWDASPPALRGLHHDLLL RKMSSSSPDLLSAALGARGRGATGGARDPGSPPPPQGDTPPSEGSAPGST  
PDSPGGAKGEP PPPVGPGEVGLLGTGREGTTGRGGS RAGYQHLTPAALLYRAAVTRSQKRGISSEEEEEGE  
DSEVELPPSQRWPQGP NMRQSLSTFSSENPSDVEEGTASEPSPSGTPEVGSTNTDERPDERSSDMCSQGSEIP  
DLPTSEVVPERETSSLPMQH QDDQGNPEDSDCDSTELDNSNSIDALPPPASLPP@MISRMIFRNYPSHNESD  
EPFHFSISRELLDRNDVVVGEMIGEGAYSIVYKGLLRNQFPVAVKIMDPSTTS AVTKAHKKTFQKEVLLLS  
MKHDNIVK FVGACIEPQLIIVTELVEGGTLQRFMHSRPGPLDLKMSLSFALDISRAMEFVHSNGIIHRDLNPR  
LLVTGDLKHVKLADFGIAREETRGGMTCEAGTSK WMAPEVYSPEPLRVGEKKEYD HKADIYSFAIVLWQL  
TNEEPFPDVPNSL FVPYLV SQGRRPILTKTPDV FVPIVESCWAQDPDARPEFKEISVMLTNLLRRMSSDSSIG  
TLPDGEAYEGEMEES ENSPLLQEHFCKVKKPKKKKKKKLVKMRFPFFKKFKVWLYNYKP@MSFFNFRKI  
KLGSEKKKKQYEHVKRDLNPEDFWEIIGELGDGA FGKVYKAQNKETSVLAAAKVIDTKSEEELEDY MVEI  
ILASCDHPNIVKLLDAFY YENNLWILIEFCAGGAVDAVMLELERPLTESQIQVVCKQTL DALNYLHDNKIIH  
DLKAGNIFTL DGDIKLADFGVSAKNTRTIQRRDSFIGTPY WMAPEVVMCETSKDRPYDYKADVWSLGITL  
EMAEIEPPHHEL NPMRVLLKIAKSEPPTLAQPSRWSSNFKDFLKKCLEKNVDARWTTSQLLQHPFVTVD SN  
PIRELIAEAKAEVTEEV EDGKEEDEEEETENSLPIASKRASSDLSIASSEEDKLSQNACILESVSEKTERS NSE  
KLNSKILNEKPTTDEPEKAVEDINEHITDAQLEAMTELHDRTAVIKENEREKRPKLENLPD TEDQETVDINS  
SEKKNIMITLETNIEHNLKSEEEKDQEKQQMFENKLIKSEEIKDTILQTVDLVSQETGEKEANIQA VDSEV  
LTKEDTQEKLGEDDKTQKDVISNTSDVIGTCEAADVAQKVDEDSAEDTQSN DGKEVVVEVGQKLINKPMVG  
EAGGTKEVP IKEIVEMNEIEEKKKK@MARTTSQLYDAVPIQSSVVL CSCPSPSMVRSQTEPSSSPGIPSGVSR  
GSTMDGT TAEARPSTNPLQQHPAQLPPQPRKKRPEDFKFGKILGEGSFSTVVLARELATSREYAIKILEKRHI  
KENKV PYPVTRERDVMSRLDHPFFVKLYFTFQDDEKLYFGLSYAKNGELLYIRKIGSFDETCTRFYTA EIVS  
LEYLHGKGIIHRDLK PENILLNEDMHIQITDFGTAKVLS PDSKQARANSFVGTAQYV SPELLTEKSACKSSD  
WALGCIIYQLVAGLP PFRAGNEYLIFQKIIKLEYDFPEKFFPKARDLVEKLLVL DATKRLGCEEMEGYGLK  
HPFFESITWENLHQQT PPKLTAYLPAMSEDDDCYGN YDNLLSQFGCMQVSSSSSSSHSLCAVDASLPQRSG  
NIEQYIHDLD TNSFELDLQFSEDEKRL LLEKQAGGNPWHQFVENNLILKMGPVDKRKGLFARRRQLLLTEG  
HLYYVDPVNKVLKGEIPWSQELRPEAKNFKTFFVHTPNRTYYLMDPSGNAHKWCRKIQEVWRQQYQSSP  
AAVQ@LTHAGWGQGWTLARTR SLLIMLGPGSNRRRPTQGERGPGSPGEPMEKYQVLYQLNPGALGVNL  
VEEMETKVKHVIKQVECMDDHYASQALEELMPLLKLRHAHISVYQELFITWNGEISSLYLCLVMEFNELSF  
EVIEDKRKAKKIIDSEWMQNVLGQVLD ALEYLHHLDIHRNLKPSNIILISSDHCKLQDLSSNVLM TDKAKW  
IRAEEDPFRKSWMAPEALNFSFSQKSDIWSLGCILDMTSCSFMDGTEAMHLRKS LRQSPGSLKAVLKTME  
KQIPDVETFRNLLPLMLQIDPSDRITIKDVVHITFLRGSFKSSCVSLTLHRQMVPASITDMLLEG NVASILEV  
QKFSGWPEVQLRAMKRLLKMPADQLGLPWPELVEVVVTTMELHDRVLDVQLCAC SLLHLLGQALVH  
PEAKAPCNQA ITSTLLSALQSHPEEEPLLVMVYSLLAITTTQESES LSEELQNAGLLEHILEHLNSSLES RDVC  
SGLGLLWALLLDGIIVNKAPLEKVPDLISQVLATYPADGEMAEASCGVFWLLSLLGCIKEQQFEQVVALLL  
SIRLCQDRALLVNNAYRGLASLVKVSELA AFKVVVQEEGSGLSLIKETYQLHRDDPEVVENVGM LLVHL  
SYEEILPELVSSSMKALLQEIKERFTSSLVSDSSAFSKPGLPPGGSPQLGCTTSGGLE@MDYYSQGT FQNIME  
KRKLKAVVDTEWMHTMLSQVLDAIEYLHKL NIVHRNLKPSNIVLVNSGYCKLQDMSSQALM THEAKWN  
RAEEDPCQKSWMAPEALKFSFSTKSDIWSLGCILDMATCSFLNDTEAMQLRKAIRHHPGSLKPILKTMEEK  
IPGTDVYYLLL PFMLHINPSDRLAIKDVMQVTFMSNSFKSSSV ALNMQRQKVPIFITDV LLEGNMANILDV  
QNFSSRPEVQLRAIN KLLTMPEDQLGLPWPTELLEEVISIIKQHGRILDILLSTCSLLRVLGQALAKDPEAEIP  
SSLIISFLMDTLRSHPN SERLVNVVYNVLAIISSQGQISEEELEEEGLFQLAQENLEHFQEDRDICLSILSLLWSL  
VDVVTV DKEPLEQLSGMVTWVLATHPEDVEIAEAGCAVLWLLSLLGCIKESQFEQVVVLLLR SIQLCPGRV  
LVNNAFRGLASLAKVSELVAFRIVVLEEGSSGLHLIQDIYKLYKDDPEVVENLCMLLAHLTSYKEILPEMES

GIKDLVQVIRGRFTSSLELISYADEILQVLEANAQPGLQEDQLEPPAGQEAPLQGEPLFRP@MEHSVPKNKL  
KLSEDSLTKQPEEVFDVLEKLGEFSYGSVFKAIHKESGQVVAIKQVPVESDLQEIIKEISIMQQCDSPYVVKY  
GSYFKNTDLWIVMEYCGAGSVSDIIRLRNKTLTETEIATVLKSTLKGLEYLHFMRKIHRDIKAGNILLNTEG  
AKLADFGVAGQLTDTMAKRNTVIGTPFWMAPEVIQEIGYNCVADIWSLGITSIEMAEGKPPYADIHPMRAI  
MIPTNPPPTFRKPEHWSDDFTDFVKKCLVKNPEQRATATQLLQHPFIVGAKPVSILRDLITEAMDMKAKRQ  
EQQRELEEDDENSEEEVEVDSHTMVKSGSESAGTMRATGTMSDGAQTMIEHGSTMLESNLGTMVINSDDDE  
EEEDLGSMRRNPTSQQIQRPSFMDYFDKQDSNKAQEGFNHNQQDPCLISKTAFPDNWKVPQDGDGDFDLKN  
DFEELQMRLTALDPMMEIEELRQRYTAKRQPILDAMDAKKRRQQNF@PTRPTRLIVSNFSQAKQKSHL  
DPQILRDQSRLAPEIITATQYKKCDEFQTGILYEMLHLPNPFDENPELKEKEYTRTDLPRIPLRSPYSWGLQQ  
ASCLLNPNPSEIRILISDAKGILQCLLWGPREDLFQIFTTSATLAQKNALLQNWLDIKRTLMIKFAEKSLDRE  
GISLEDWLCAQYLAFATTDLSYIVKILQYR@MQNKENREPRVQQTPSAGVGPLRVEMNPDTHAVSGPGR  
PVKSNSKVLSIDDFDIGRPLGKGKFGNVYLARERKLKVIALKVLFKSQMVKEGVEHQLRREIEIQSHLRHP  
ILRFYNYFHDDTRVFLILEYAPRGEMYKELQRYGRFDDQRTATYMEEVSDALQYCHEKKVIHRDIKPENLL  
GYRGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVGNPPFETASHA  
TYKRITKVDLQFPKLVSEGARDLISKLLRHSPSMRLPLRSVMEHPWVKANSRRVLPPVCSSEPH

Q5b.

6978799 0

45768856 509

45768786 943

27923854 1367

45768720 1765

45768758 2292

45219906 2639

18202599 3311

4835224 4200

40787731 4571

18202068 5189

34191428 5749

29747774 6450

28856169 7038

20071571 7531

45709347 7734

Q6.

18202068

Q7.

The assignment took me roughly 8 hours to complete.

I worked alone on this assignment and did not reach out for help.