

Salmon Age-Sex-Length Data Transfer Protocol

Variations of format, coding, and definitions have been used among regions in collecting Alaska's detailed salmon Age-Sex-Length data. Differences have compounded over time, as staffing and program requirements changed. This protocol is defined to facilitate the transmission of Alaska's historic and future salmon Age-Sex-Length biological data into a central repository. It provides a clear specification that explains the nature of data to the users, and a well-defined transmission mechanism for populating the repository.

A. DRAFT Data Submission Specification v2.1

Max Chars refers to the maximum number of characters a field's value may contain. **Reqd** indicates whether the field must contain a value (nulls ARE NOT accepted when Reqd=Yes). **Data Type** references how the column should be defined in a relational database. **Validation** gives specific rules that the submission must fully meet in order to be accepted into the repository.

Division of Commercial Fisheries ASL Repository Data Submission Specification						
Ref	Column Name (Alternate Name)	Max Chars	Reqd	Data Type	Description	Validation
1	Sample_ID	23		Character	Value generated by the reporting region, which may be used to associate specific records in the original regional data as belonging to a particular sampling event. Uniqueness is desired, but not mandatory.	
2	Region_ID	1	Yes	Character	Commercial Fisheries region of collection	Must be '1', '2', '3' or '4'
3	Sample_Year	4	Yes	Character	4-digit year in which sampling event occurred	Must be between 1930 and the current calendar year
4	Management_area	3		Character	Geographic area spanning a number of districts. Typically defined at the region level for their areas of interest.	
5	Tix_management_area	1		Character	Management Area code formally defined by the fish ticket system.	If present, must match an existing code in Appendix A
6	District	3		Character	Three character district of observation	If present, must match a current or

Division of Commercial Fisheries ASL Repository Data Submission Specification

Ref	Column Name (Alternate Name)	Max Chars	Reqd	Data Type	Description	Validation
						historic CF fish ticket district
7	Subdistrict	3		Character	ID for subset of district sampled, if any	If present, must be exactly two digits
8	Stream	50		Character	ID for anadromous waters catalog stream sampled R 3 always 0	
9	Location	3		Character	R 1: port codes R 2: stream location R 3: stream location, some fish tickets R 4: some stream locations	
10	Project	2		Character	“Fishery Type” Legacy code. Typical values are in Appendix I	
11	Sample_Day	2		Character	Day this sample was taken, or began to be taken.	If present, must be 1 or 2 digits in range 1 through 31
12	Sample_Month	2		Character	Month number when this sample was started	If present, must be 1 or 2 digits in range 1 through 12
13	Sample_Date	10		Date mm/dd/yyyy	Single date sample was taken	If present, must be a valid date.
14	Gear	2		Character	Type of collection gear. Typical values are in Appendix H.	
15	Harvest_Code	2		Character	Type of commercial fishery sampled	If present, must match a current or historic Harvest Code in Appendix B
16	Mesh	5		Character	Net mesh size Converted to inches None for R1 and R4	
17	Length_Type (Measurement Type)	2		Character	Codes indicating type of length measurement	If present, must match a mark-sense length code in Appendix C
18	Number_Scales	1		Number	Number of scales per fish	If present, must be a valid whole number with no punctuation

Division of Commercial Fisheries ASL Repository Data Submission Specification

Ref	Column Name (Alternate Name)	Max Chars	Reqd	Data Type	Description	Validation
19	Number_Cards	1		Number	Number of gum cards taken for a particular mark-sense form referenced in field 20	If present, must be a valid whole number with no punctuation
20	Form_Number	15		Character	Mark-sense data form sequence number	If present, must be digits
21	Species	3	Yes	Character	Salmon species code	Must match a Species Code in Appendix D
22	Stage	1	Yes	Character	Indicates Juvenile, adult, etc.	Must match a stage code in Appendix E
23	Batch_Number (Data Link)	24		Character	Region specific – track uploading of data or data Source	
24	Stat_week	2		Character	Statistical week	If present, must be 1 or 2 characters representing a number between 1 and 54
25	Period	2		Character	Openings (e.g. R3)	
26	Comments	50		Character		
27	Specimen_ID	4		Character	Value generated by the submitter which identifies a particular observation in a sample. Typically, they range 1 through n.	
28	Card_Number	3		Character	Sequence number of gum card in a particular collection	
29	Fish_Number	5		Character		
30	Sex	1		Character	Male, female, indeterminate	If present, must match an existing sex code in Appendix F
31	Length	4		Number	Length of fish in millimeters	If present, must be a valid whole number
32	FW_Age	1		Character	Freshwater age using European method	If present, must be a digit
33	SW_Age	1		Character	Saltwater age using European method	If present, must be a digit
34	Age_Error_code	10		Character	String of one or more digits indicating problems in aging.	If present, must be composed of digits 0 through 9

Division of Commercial Fisheries ASL Repository Data Submission Specification

Ref	Column Name (Alternate Name)	Max Chars	Reqd	Data Type	Description	Validation
					Regional use varies somewhat. Typical values are in Appendix G.	
35	Weight	7		Number	Weight in grams to the nearest tenth of a gram	If present, must be a number in the range 0.1 through 99999.9
36	CWT_Head_Number (Strap Tag, Cinch Strap)	6		Character	Six digit strap tag number identifying a head collected for the coded wire tag lab	If present, must be all digits
37	Mark_Recapture_Tag	10		Character	Up to 10 characters from a tag used in mark-recapture programs: disk tag, spaghetti tag, etc.	
38	DNA_VIAL_CODE	25		Character	Up to 25 characters used to identify DNA specimen collected for genetics lab and placed in a Vial which was assigned a DNA_VIAL_CODE	
39	Otolith_BP_Number (BP Coordination #, Brain Parasite #)	8		Character	Up to 8 characters used to identify a specimen collected for otolith lab	
40	Image_Name	20		Character	File name for digitized scale image	
41	Format	6	Yes	Character	The version of the transfer specification used to build the file. Every row in the file must have this same value.	Must be "CF-2.1" 2.0 will continue to be accepted based on 2.0 specification rules *
42	Otolith_Tray_Cell	6		Character	Six digit key to Mark Lab otolith recoveries: 3 digit tray number concatenated with 3 digit cell number	If present, must be exactly 6 (i.e., zero filled)
43	Quadrant	2		Character	Two character quadrant code from Coded Wire Tag lab	Must match quadrant in Appendix J
44	ADFG_Number	5		Character	Up to 5 digit ADF&G vessel number from CFEC	If present, must be composed of digits 0 through 9

Division of Commercial Fisheries ASL Repository Data Submission Specification

Ref	Column Name (Alternate Name)	Max Chars	Reqd	Data Type	Description	Validation
45	Processor_code	5		Character	Five character “intent to operate” processor code from Comm Fish HQ	If present, must be one alpha followed by 4 digits
46	DNA_TRAY_CODE	25		Character	Up to 25 characters used to identify DNA specimen collected for genetics lab and placed in a Tray which was assigned a DNA_TRAY_CODE	
47	SSID	13		Number	Standard Specimen Identifier	If present must be unique
48	DNA_TRAY_WELL_CODE	3		Number	Identifies the Column and Row within a DNA Tray, referenced in 46	
49	DNA_TRAY_WELL_POSITION	3		Character	Identifies the Column and Row within a DNA Tray in which a specimen resides which may be necessary to determine the orientation of a tray as it was filled with specimens	

[Items in GREEN were specifically defined for the CF-2.1 revision]

* Version 2.0 will continue to be accepted based on specification 2.0 rules when the value of FORMAT is 2.0 affecting specifically:
 DNA_NUMBER will be an accepted column name, however; the value will be placed in the DNA_VIAL_NUMBER column
 DNA_TRAY_CODE, SSID, DNA_TRAY_WELL_CODE, and DNA_TRAY_WELL_POSITION will not be required columns
 Submitting agencies are encouraged to submit based on version 2.1 rules whenever possible.

B. Transmission Mechanism

The initial transmission mechanism addresses populating the repository with historic data. Once a solid set of experience is obtained in this process, the mechanism for collecting current season data will be addressed in detail.

Data Characteristics:

1. All submitted data must be presented in comma-separated value (CSV) files using a Windows-compatible character set, preferably a version of printable ASCII.
2. All files must contain only newline-delimited records. That is, there must be one record per line of the file.
3. Any field whose value contains an embedded comma (,) must be surrounded in double-quotes (“”). It is permissible for all fields to be reported with surrounding double-quotes. Any double-quote delimiters will be stripped from the fields as they are stored in the database.
4. No double-quotes are allowed as data values of any data field. The double-quote is sequestered for exclusive use as a field delimiter.
5. The first record in the file must contain Column Names as they are defined in the specification. This serves as inline documentation. The first row of a file will always be skipped when the repository is loaded. But the header will make any submitted file readily identifiable to staff managing the repository, regardless of the file’s name.
6. All fields which do not contain a data value are considered NULL. The fields for which data are absent must be denoted in the file using two consecutive commas (,).

File Scope:

1. For purposes of reporting historic data, a report file shall be all the data for a particular REGION_ID in a particular SAMPLE_YEAR. Those values must be constant in every record of a particular file.
2. Each file submission will be validated according to the mandatory rules in the Specification. Any rows having invalid data will be reported back to the originator. The originator will resubmit the complete file with errors corrected.
3. As soon as a file is determined to meet the validation rules of the specification, it will be loaded into the repository. Before inserting the file contents into the database, all existing data for the region and year will be destroyed. This is necessary to prevent multiple copies of records from accumulating in the repository – records are not required to have unique keys in the repository and cannot be readily deleted or updated on an individual basis.
4. Files may be submitted to the repository by uploading them to <ftp://FTP.TAGLAB.ORG/UPLOAD/ASL/REGIONx> where ‘x’ is a digit 1, 2, 3, or 4 representing the reporting region.
5. Because there is adequate bandwidth and file space for the repository, submissions should not be compressed.

Appendix A – Standard Tix Management Area Codes

<u>CODE</u>	<u>DESCRIPTION</u>
9	CANADA
A	JUNEAU/YAKUTAT
B	KETCHIKAN/CRAIG
C	PETERSBURG/WRANGELL
D	SITKA/PELICAN
E	PRINCE WILLIAM SOUND
F	EEZ
H	COOK INLET
K	KODIAK
L	CHIGNIK
M	ALASKA PENINSULA
O	DUTCH HARBOR
Q	BERING SEA
R	ADAK/WEST ALEUTIANS
S	SOUTHEAST INSIDE (1990-1998)
T	BRISTOL BAY
W	KUSKOKWIM
X	KOTZEBUE
Y	YUKON
Z	NORTON SOUND

COMMERCIAL FISHERIES

Appendix B – Standard Harvest Codes

<u>CODE</u>	<u>DESCRIPTION</u>
11	TRADITIONAL
12	TERMINAL AREA
13	EXPERIMENTAL AREA
14	EXPERIMENTAL GEAR
17	M-I-C
18	CONFISCATED
21	PNP FISH
22	PNP CARCASSES
23	STATE FISH
24	STATE CARCASSES
25	FEDERAL FISH
26	FEDERAL CARCASSES
27	PNP DONATED
28	PNP DISCARDED
31	DERBY
33	DISCARDED
34	OILED WASTE
35	EDUCATIONAL
36	COMMERCIAL DONATED
41	TEST RUN ASSESSMENT
42	TEST SPECIAL STUDY
43	TEST STOCK ASSESSMENT

COMMERCIAL FISHERIES

Appendix C – Standard Length Type Codes

<u>CODE</u>	<u>DESCRIPTION</u>
00	LENGTH NOT TAKEN
01	TIP OF SNOUT TO FORK OF TAIL
02	MID-EYE TO FORK OF TAIL
03	POST ORBIT TO FORK OF TAIL
04	MID-EYE TO HYPURAL PLATE
05	POST ORBIT TO HYPURAL PLATE
06	TIP OF SNOUT TO TIP OF TAIL
07	CLEITHRAL ARCH TO TIP OF TAIL
08	CALCULATED FORK LENGTH

COMMERCIAL FISHERIES

Appendix D – Standard Species Codes

<u>CODE</u>	<u>DESCRIPTION</u>
410	CHINOOK
420	SOCKEYE
430	COHO
440	PINK
450	CHUM
470	CUTTHROAT
540	STEELHEAD
666	ATLANTIC

COMMERCIAL FISHERIES

Appendix E – Standard Stage Codes

<u>CODE</u>	<u>DESCRIPTION</u>
A	ADULT
E	EMERGENT FRY
F	FED FRY
G	FINGERLING
I	IMMATURE
J	JUVENILE
P	PRESMOLT
S	SMOLT

COMMERCIAL FISHERIES

Appendix F – Standard Sex Codes

<u>CODE</u>	<u>DESCRIPTION</u>
1	MALE
2	FEMALE
3	EXAMINED BUT DID NOT IDENTIFY

COMMERCIAL FISHERIES

Appendix G – Typical Age Error Codes

<u>CODE</u>	<u>DESCRIPTION</u>
1	OTOLITH
2	INVERTED
3	REGENERATED
4	ILLEGIBLE
5	MISSING
6	REABSORBED
7	WRONG SPECIES
8	NOT PREFERRED

COMMERCIAL FISHERIES

Appendix H – Typical Gear Codes

<u>CODE</u>	<u>DESCRIPTION</u>
00	TRAP
01	PURSE SEINE
02	BEACH SEINE
03	DRIFT GILLNET
04	SET GILLNET
05	HAND TROLL
06	LONG LINE
07	OTTER TRAWL
08	FISHWHEEL
09	POTS
10	SPORT HOOK AND LINE
11	HERRING PURSE SEINE
12	HANDPICKED
13	DIP NET
14	WEIR
15	POWER TROLL
17	BEAM TRAWL
18	SHOVEL
19	WEIR
90	TRAP
91	POTS

COMMERCIAL FISHERIES

Appendix I – Typical Project Codes

<u>CODE</u>	<u>DESCRIPTION</u>
1	COMMERCIAL HARVEST
2	SUBSISTENCE HARVEST
3	ESCAPEMENT (TOWER, WEIR, SONAR SITE, ETC.)
4	ESCAPEMENT – SPAWNING GROUNDS
5	TEST FISHING
6	SPORT CATCH (MARINE)
7	SPORT CATCH (FRESHWATER)

COMMERCIAL FISHERIES

Appendix J – Quadrant Codes

<u>QUADRANT</u>	<u>QUADRANT DESCRIPTION</u>	<u>REGION</u>
NE	SE - NORTHEAST	1
NW	SE - NORTHWEST	1
SE	SE - SOUTHEAST	1
SW	SE - SOUTHWEST	1
BB	BRISTOL BAY	2
BS	BERING SEA	2
LC	LOWER COOK INLET	2
PW	PRINCE WILLIAM SOUND	2
UC	UPPER COOK INLET	2
KO	KOTZEBUE	3
KU	KUSKOKWIM	3
NS	NORTON SOUND	3
YU	YUKON	3
AL	ALEUTIANS	4
CH	CHIGNIK	4
DH	DUTCH HARBOR	4
KD	KODIAK	4
PE	PENINSULA	4

COMMERCIAL FISHERIES